

# On the Many-Help-One Problem with Independently Degraded Helpers

Albrecht Wolf, Diana Cristina González, Meik Dörpinghaus, *Member, IEEE*,  
José Cândido Silveira Santos Filho, *Member, IEEE*, and Gerhard Fettweis, *Fellow, IEEE*

## Abstract

This work provides new results for the separated encoding of correlated discrete memoryless sources. We address the scenario where an arbitrary number of auxiliary sources (a.k.a. helpers) are supposed to help decode a single primary source errorless. The rate-distortion analysis of such a system establishes the so-called many-help-one problem. We focus on the case in which the auxiliary sources are conditionally independent, given the primary source. We provide an inner and an outer bound of the rate-distortion region for what we call the strong many-help-one problem, where the auxiliary sources must be also recovered under certain distortion constraints. Furthermore, based on known results from the CEO problem we provide the admissible rate region for what we call the weak many-help-one problem, where the auxiliary sources do not need to be recovered serving merely as side information to help recover the primary source. Both scenarios find important application to emerging cooperative techniques in which a direct-link communication is assisted by multiple lossy relaying-link transmissions. Motivated by this application, we specialize the referred scenarios to binary sources and the Hamming distortion measure. Remarkably, we show that for the weak many-help-one problem the independent decoding of the auxiliary sources can achieve optimal performance.

## Index Terms

Admissible rate region, correlated sources, distributed source coding, many-help-one problem, rate-distortion region.

## I. INTRODUCTION

Multiterminal source coding has a rich history. Slepian and Wolf [1] were the first to characterise the problem of distributed encoding of multiple correlated sources. In their seminal paper, the admissible rate region for the lossless distributed encoding of two correlated sources was derived. A simple proof of the Slepian-Wolf result with extension to an arbitrary number of correlated sources was presented by Cover [2]. Wyner [3] and Ahlswede and Körner [4] considered a different problem, in which an auxiliary random variable (i.e., side information) was introduced to expand the rate region of a lossless single-source coding problem. In that setup coded (or partial) side information was available at the decoder. Wyner and Ziv [5] presented a generalization to lossy single-source coding with uncoded side information. This was the first characterization of a multiterminal rate-distortion function. Berger [6] and Tung [7] extended the Slepian-Wolf problem to the lossy distributed encoding of an arbitrary number of correlated sources. In those two works, inner and outer bounds were presented for the multiterminal rate-distortion region, which do not coincide in general. Berger and Yeung [8] addressed a variant of the Wyner-Ziv problem, namely, the distributed encoding of two sources, one of which allows for lossy compression. Over the years, a significant progress has been made in the area of multiterminal source coding, including general frameworks for lossless compression [9], [10], results for Gaussian sources in various contexts [11], [12], and Wyner-Ziv networks containing an arbitrary number of sources [13]. Yet, an exact solution to the multiterminal rate-distortion region remains open, even for the simplest scenario considered in [6].

The idea that a decoder wishes to reproduce a primary source with the help of an auxiliary source, introduced by Wyner, Ahlswede, and Körner, can be intuitively extended to an arbitrary number of auxiliary sources. Finding of the admissible rate region of such a system defines the so-called many-help-one problem. This problem has been recognized as a highly challenging one and only a few particular solutions are known to date. Körner and Marton [14] addressed a two-help-one problem where the primary source is a modulo-two sum of correlated binary auxiliary sources. Gelfand and Pinsker [15] determined the admissible rate region when the auxiliary sources are discrete, memoryless, and conditionally independent if the primary source is given. Motivated by the Gelfand-Pinsker result, Oohama [16] determined the rate-distortion region for the same setup but Gaussian sources. Tavildar [12] derived the rate-distortion region for Gaussian memoryless sources with a correlation model following a tree-structure.

This work contributes new results to the many-help-one problem for lossless compression of the primary source. As in

The work of A. Wolf was supported by the FP7 project ICT-619555 RESCUE (Links-on-the-fly Technology for Robust, Efficient and Smart Communication in Unpredictable Environments) which is partly funded by the European Union.

The work of M. Dörpinghaus was supported in part by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the cluster of excellence EXC 1056 “Center for Advancing Electronics Dresden (cfaed)” and by the DFG within the Collaborative Research Center SFB 912 “Highly Adaptive Energy-Efficient Computing (HAEC)”.

A. Wolf, M. Dörpinghaus, and G. Fettweis are with the Vodafone Chair Mobile Communications Systems, Technische Universität Dresden, Dresden, Germany, E-mails: {albrecht.wolf, meik.doerpinghaus, fettweis}@tu-dresden.de.

D. C. González and J. C. S. Santos Filho are with the Department of Communications, School of Electrical and Computer Engineering, University of Campinas (UNICAMP), Campinas-SP, Brazil, E-mails: {dianigon, candido}@decom.fee.unicamp.br.

[15] and [16], we consider that the auxiliary sources (a.k.a. helpers) are conditionally independent given the primary source. This is equivalent to assume that the auxiliary sources represent independently degraded versions of the primary source, thus finding direct application to emerging cooperative communication schemes with multiple lossy relaying routes from source to destination. On the other hand, unlike [15] and [16], we focus on the binary case, to which our major contributions refer. These contributions are detailed in Section I-F. We introduce two variations of the many-help-one problem, denoted as follows: (i) the strong many-help-one problem, where not only the primary source must be recovered errorless, but also the auxiliary sources themselves must be recovered under any given set of distortion constraints, and (ii) the weak many-help-one problem, where only the primary source must be recovered, the auxiliary sources serving merely as side information.

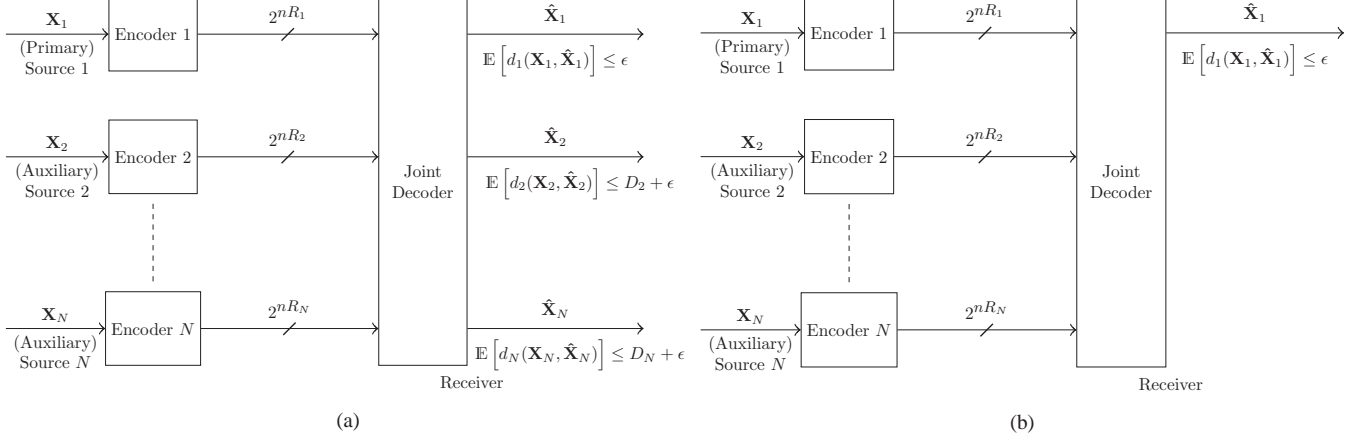


Fig. 1: (a) The strong many-help-one problem; (b) the weak many-help-one problem.

#### A. The Strong Many-Help-One Problem

The strong many-help-one problem is presented in Fig. 1a. It has  $N$  discrete memoryless sources, denoted as  $\{(X_{1,k}, X_{2,k}, \dots, X_{N,k})\}_{k=1}^{\infty}$ , with the  $n$ -sample sequence of the  $i$ th source being represented in vector form as  $\mathbf{X}_i = [X_{i,1}, X_{i,2}, \dots, X_{i,n}]$ ,  $i = 1, \dots, N$ . When appropriate, for simplicity, we shall drop the temporal index of the sequences, denoting the sources merely as  $X_1, X_2, \dots, X_N$ . By assumption,  $X_1, X_2, \dots, X_N$  are mutually correlated random variables, taking values in finite sets  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_N$ , respectively, and are distributed according to a fixed and known probability mass function (pmf)  $p(x_1, x_2, \dots, x_N)$ . The encoders generate binary sequences with rates  $R_1, R_2, \dots, R_N$  bits per input symbol. The decoder output is a sequence of  $N$ -tuples  $\{(\hat{X}_{1,k}, \hat{X}_{2,k}, \dots, \hat{X}_{N,k})\}_{k=1}^{\infty}$  whose components take values from finite reproduction alphabets  $\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2, \dots, \hat{\mathcal{X}}_N$ . The encoding is done in blocks of length  $n$  and the distortion constraint is given for the primary source  $X_1$  by

$$\mathbb{E} \left[ \frac{1}{n} \sum_{k=1}^n d_1(X_{1,k}, \hat{X}_{1,k}) \right] \leq \epsilon \quad (1)$$

and for the auxiliary sources  $X_i, i = 2, \dots, N$ , by

$$\mathbb{E} \left[ \frac{1}{n} \sum_{k=1}^n d_i(X_{i,k}, \hat{X}_{i,k}) \right] \leq D_i + \epsilon, \quad (2)$$

where  $d_i(x_i, \hat{x}_i) \geq 0, x_i \in \mathcal{X}_i, \hat{x}_i \in \hat{\mathcal{X}}_i$ , is a given distortion function which can differ for each source. Moreover,  $\mathbb{E}[\cdot]$  denotes the expectation operator and  $\epsilon$  is a small positive number. The strong many-help-one problem can be seen as a discrete multiterminal source coding with  $N - 1$  distortion constraints assigned to the auxiliary sources.

#### B. The Weak Many-Help-One Problem

The weak many-help-one problem is presented in Fig. 1b. It is similar to the strong many-help-one problem, except that the decoder output is only the sequence  $\{\hat{X}_{1,k}\}_{k=1}^{\infty}$  whose components take values from the finite reproduction alphabet  $\hat{\mathcal{X}}_1$ . The encoding is done in blocks of length  $n$  and the only distortion constraint is given for the primary source  $X_1$  by

$$\mathbb{E} \left[ \frac{1}{n} \sum_{k=1}^n d_1(X_{1,k}, \hat{X}_{1,k}) \right] \leq \epsilon. \quad (3)$$

The auxiliary sources  $X_i, i = 2, \dots, N$ , are not reconstructed at the receiver.

### C. Independently Degraded Helpers

In this work, following [15] and [16], we consider that the auxiliary sources (a.k.a. helpers) are conditionally independent given the primary source. Hereafter, this scenario is referred to as the CI condition. Thus the joint probability distribution of  $X_1, X_2, \dots, X_N$  satisfies

$$p(x_1, x_2, \dots, x_N) = p(x_1) \prod_{i=2}^N p(x_i | x_1). \quad (4)$$

As mentioned before, the CI condition finds important application to emerging cooperative communication techniques based on lossy relaying links.

### D. Problem Statement

We define the rate-distortion region  $\mathcal{R}(D_2, \dots, D_N)$  for the strong many-help-one problem and the admissible rate region  $\mathcal{R}$  for the weak many-help-one problem as the set of rate  $N$ -tuples  $R_1, R_2, \dots, R_N$  for which the systems in Fig. 1a or Fig. 1b satisfy the average distortions constraints in (1) and (2), or in (3), as required, for  $n \rightarrow \infty$ , respectively. The essential problem is to characterize the set of minimum rate  $N$ -tuples at which the encoders can communicate with the decoder while still conveying enough information to satisfy the distortion constraints on the reconstruction.

### E. Known Results

*a) Strong many-help-one problem:* Berger [6] and Tung [7] studied the problem of  $N$  memoryless discrete sources with arbitrary distortion constraints  $D_1, D_2, \dots, D_N$ . In [6, Theorem 6.1 and Theorem 6.2], an inner bound, i.e., an achievable rate-distortion region,  $\mathcal{R}_a(\cdot) \subseteq \mathcal{R}(\cdot)$ , and an outer bound  $\mathcal{R}_c(\cdot) \supseteq \mathcal{R}(\cdot)$  were provided, respectively. Eventually, the inner and the outer bound do not coincide and thus the rate-distortion region remains unknown for the referred problem.

Gastpar [13] extended the Berger-Tung problem by providing *uncoded* side information at the decoder. Alternatively, Gastpar's work can be seen as an extension of the Wyner-Ziv problem [5] to  $N$  sources. In [13, Theorem 2 and Theorem 3], an inner and an outer bound were provided for the rate-distortion region, respectively. However, similarly to the Berger-Tung problem, those bounds do not coincide. This is not surprising, indeed, as the analysis in [13] is an extension of the Berger-Tung problem. In [13, Section V], a partial converse was presented, with the assumption that all the sources are conditionally independent given the side information. It was shown in [13, Theorem 6] that in this case the inner and the outer bound coincide.

Berger and Yeung [8] investigated the Berger-Tung problem when specialized to two memoryless discrete sources with distortion constraints  $D_1 = 0$  and  $D_2$  is arbitrary. Note that this setup corresponds to the strong one-helps-one problem. In [8, Section III and Section IV], inner and outer bound were provided, respectively. Once again, the two bounds do not coincide. Additionally, in [8, Section VI], binary sources were considered which are interconnected via a binary symmetric channel (BSC) and the distortion was measured by the Hamming distance. It was then shown that, depending on  $D_2$ , the inner and the outer bound can coincide.

Bounds on the rate-distortion region of the strong many-help-one problem in Fig. 1a can be derived by specializing the results in [6], [7], and [13], i.e., by setting a zero distortion constraint for source one, that is,  $D_1 = 0$ . However, no publication so far addressed the problem of conditionally independent auxiliary sources given the primary source. Moreover, bounds with binary sources and Hamming distortion measure are considered in neither [6], [7], nor [13]. Berger and Yeung [8] did address binary sources and Hamming distortion measure, but only for two sources.

*b) Weak many-help-one problem:* Gelfand and Pinsker [15] determined the admissible rate region for a discrete lossless version of the so-called CEO (Chief Executive Officer) problem, under the same CI condition assumed herein. Note that the weak many-help-one problem is a special case of the lossless CEO problem. However, Gelfand and Pinsker did not address binary sources and the Hamming distortion measure, for which we shall obtain insightful particular results.

### F. Contributions of this Work

The contributions of this paper include the following:

- 1) In Section II-A and Section II-B, we provide bounds on the rate-distortion region  $\mathcal{R}(D_2, \dots, D_N)$  for the strong many-help-one problem. More specifically, an inner bound  $\mathcal{R}_a(D_2, \dots, D_N) \subseteq \mathcal{R}(D_2, \dots, D_N)$  and an outer bound  $\mathcal{R}_c(D_2, \dots, D_N) \supseteq \mathcal{R}(D_2, \dots, D_N)$  are given based on the results in [13]. As in [13], our bounds do not coincide, which is not surprising, since we address a special case of the Berger-Tung problem. Our bounds generalize those presented in [17] for two sources.
- 2) We investigate the CI case of the strong many-help-one problem, where the auxiliary sources are independent while conditioned on the primary source in Section II-D. The inner and the outer bound do not coincide for this case. This

result differs from that presented in [13, Section V], where the inner and the outer bound coincide if all the sources are conditionally independent given *uncoded* side information. Note, however, that the conditional independence in [13] also differs from the one we consider, as in the former the conditioning variable is not the primary source, but some side information.

- 3) In Section II-E, we derive single-letter expressions for the inner and the outer bound of the rate-distortion region in 2) with binary sources and Hamming distortion measure. We show, that depending on the distortion constraints  $D_2, \dots, D_N$ , the inner and the outer bound can coincide.
- 4) In Section III-A, based on the results from Gelfand and Pinsker [15], we derive the admissible rate region of the weak many-help-one problem. Furthermore, we derive single-letter expressions for the admissible rate region with binary sources, CI condition, and Hamming distortion measure in Section III-B. We show that independent decoders for the auxiliary sources are optimal, which is a remarkable result having operational meaning.

### G. Notation and Terminology

The upper- and lowercase letters are used to denote random variables and their realizations, respectively. The alphabet set of a random variable  $X$  with realization  $x$  is denoted by  $\mathcal{X}$ , and its cardinality, by  $|\mathcal{X}|$ . The pmf of the random variable  $X$  is denoted by  $p_X(x)$ , or simply  $p(x)$  when this does not create any confusion. Also,  $\mathbf{X}$  and  $\mathbf{x}$  represent vectors containing a temporal sequence of  $X$  and  $x$ , respectively. All random vectors have length  $n$  (block length). We use  $k$  to denote the time index and  $i$  to denote a source index. We define  $\mathbf{A}_S \triangleq \{\mathbf{A}_i | i \in S\}$  as an indexed series of random vectors,  $\mathbf{A}_S(b_i) \triangleq \{\mathbf{A}_i(b_i) | i \in S\}$  as an indexed series of random time vectors with variable  $b_i$  and  $A_S \triangleq \{A_i | i \in S\}$  as an indexed series of random variables. In general, a vector  $\mathbf{A}$  contains elements  $a_{(\cdot)}$ , as in  $\mathbf{A} = \{a_1, a_2, \dots, a_{|\mathbf{A}|}\}$ . We define two particular sets:  $\mathcal{N} = \{1, 2, \dots, N\}$  and  $\mathcal{L} = \{2, 3, \dots, N\}$ . The binary entropy function is denoted as  $h(p) = -p \log(p) - (1-p) \log(1-p)$ . The probability of an event  $\mathcal{E}$  is denoted as  $\Pr[\mathcal{E}]$ . The binary convolution is defined as  $a_1 * a_2 = a_1(1-a_2) + (1-a_1)a_2$ . The multivariate binary convolution is defined as  $a_1 * \dots * a_N = a_1 * (\dots * (a_{N-1} * a_N) \dots)$ , which is a *cascaded* binary convolution.

Moreover, we shall make use of the concept of strong typicality as introduced in [6]. We use the definition as in [18, p. 326]. Let  $N(a|\mathbf{x})$  be the number of occurrences of a symbol  $a$  in the sequence  $\mathbf{x}$ . The sequence  $\mathbf{x}$  is said to be  $\epsilon$ -strong typical with respect to a pmf  $p_X(x)$  if,  $\forall a \in \mathcal{X}$ ,  $|1/n N(a|\mathbf{x}) - p_X(a)| < \epsilon/|\mathcal{X}|$ . For a given pmf  $p_X(x)$ , the set of  $\epsilon$ -strong typical sequences will be denoted by  $\mathcal{T}_\epsilon^{*(n)}(X)$ , or simply  $\mathcal{T}_\epsilon^{*(n)}$ .

## II. THE STRONG MANY-HELP-ONE PROBLEM

In this section we investigate the strong many-help-one problem depicted in Fig. 1a and provide an inner and an outer bound for the rate-distortion region.

### A. An Inner Bound

An inner bound can be obtained based on the coding scheme introduced by Berger in [6]. In this scheme, the sequence  $\mathbf{X}_i$  is encoded in two stages. First, a code is used with a suitable quantization vector referred to as codeword; second, a binning operation for the codeword indices is applied. Herein, different from [6], encoder one has to meet a zero distortion constraint, i.e.,  $D_1 = 0$ , and thus the quantization is tightly restricted, if possible at all. All encoders use a joint codebook, but operate independently. Each encoder applies a binning operation with respect to the codeword index. Based on the bin index, the decoder aims at the perfect reconstruction of codewords. This requires all reconstructed codewords to be jointly typical.

*Theorem 1:*  $\mathcal{R}_a(D_{\mathcal{L}}) \subseteq \mathcal{R}(D_{\mathcal{L}})$ , where  $\mathcal{R}_a(D_{\mathcal{L}})$  is the set of all rate  $N$ -tuples  $R_N$  such that there exists a  $N-1$ -tuple  $U_{\mathcal{L}}$  of discrete random variables with

$$p(u_{\mathcal{L}}, x_N) = p(x_N) \prod_{i=2}^N p(u_i | x_i) \quad (5)$$

for which the following conditions are satisfied:

$$\sum_{i \in S} R_i \geq \begin{cases} H(X_1 | U_{S^c}) + I(X_{S \setminus \{1\}}; U_{S \setminus \{1\}} | X_1, U_{S^c}) & \text{if } 1 \in S, \\ I(X_S; U_S | X_1, U_{S^c \setminus \{1\}}) & \text{otherwise,} \end{cases} \quad (6a)$$

$$\quad (6b)$$

$\forall S \subseteq \mathcal{N}$ , with  $S^c$  being the complement of  $S$ ; and for which there exist functions  $g_i(\cdot), \forall i \in \mathcal{N}$ , such that for

$$\mathbb{E}[d_1(X_1, g_1(X_1, U_{\mathcal{L}}))] \leq \epsilon \quad (7)$$

and  $\forall i \in \mathcal{L}$

$$\mathbb{E}[d_i(X_i, g_i(X_1, U_{\mathcal{L}}))] \leq D_i + \epsilon. \quad (8)$$

*Proof:* See Appendix A. ■

### B. An Outer Bound

We now present a general outer bound  $\mathcal{R}_c(D_{\mathcal{L}})$  which contains the desired rate-distortion region  $\mathcal{R}(D_{\mathcal{L}})$ , i.e.,  $\mathcal{R}_c(D_{\mathcal{L}}) \supseteq \mathcal{R}(D_{\mathcal{L}})$ . Our derivation follows the standard outer-bound arguments given in [6, Theorem 6.2].

*Theorem 2:*  $\mathcal{R}_c(D_{\mathcal{L}}) \supseteq \mathcal{R}(D_{\mathcal{L}})$ , where  $\mathcal{R}_c(D_{\mathcal{L}})$  is the set of all rate  $N$ -tuples  $R_{\mathcal{N}}$  such that there exists a  $N - 1$ -tuple  $U_{\mathcal{L}}$  of discrete random variables whose joint pmf  $p(u_{\mathcal{L}}, x_{\mathcal{N}})$  must satisfy the marginal constraints

$$\sum_{u_j, j \in \mathcal{L} \setminus \{i\}} p(u_{\mathcal{L}}, x_{\mathcal{N}}) = p(x_{\mathcal{N}})p(u_i|x_i), \quad (9)$$

$\forall i \in \mathcal{L}$ , for which the following conditions are satisfied:

$$\sum_{i \in \mathcal{S}} R_i \geq \begin{cases} H(X_1|U_{\mathcal{S}^c}) + I(X_{\mathcal{L}}; U_{\mathcal{S} \setminus \{1\}}|X_1, U_{\mathcal{S}^c}) & \text{if } 1 \in \mathcal{S}, \\ I(X_{\mathcal{L}}; U_{\mathcal{S}}|X_1, U_{\mathcal{S}^c \setminus \{1\}}) & \text{otherwise,} \end{cases} \quad (10a)$$

$$(10b)$$

$\forall \mathcal{S} \subseteq \mathcal{N}$ ; and for which there exist functions  $g_i(\cdot)$ ,  $\forall i \in \mathcal{N}$ , such that

$$\mathbb{E}[d_1(X_1, g_1(X_1, U_{\mathcal{L}}))] \leq \epsilon \quad (11)$$

and  $\forall i \in \mathcal{L}$

$$\mathbb{E}[d_i(X_i, g_i(X_1, U_{\mathcal{L}}))] \leq D_i + \epsilon. \quad (12)$$

*Proof:* See Appendix B. ■

### C. Discussion

Since  $\mathcal{R}_a$  and  $\mathcal{R}_c$  do not coincide in general, the exact rate-distortion region is unknown for the strong many-help-one problem in Fig. 1a. The most important distinction between  $\mathcal{R}_a$  in Theorem 1 and  $\mathcal{R}_c$  in Theorem 2 is that one between the condition on pmf  $p(x_{\mathcal{N}}, u_{\mathcal{L}})$  in (5), i.e.,  $U_i \rightarrow X_i \rightarrow X_j \rightarrow U_j$ , and the condition on pmf  $p(x_{\mathcal{N}}, u_{\mathcal{L}})$  in (9), i.e.,  $U_i \rightarrow X_i \rightarrow X_j, X_i \rightarrow X_j \rightarrow U_j, \forall i \in \mathcal{L}, \forall j \in \mathcal{L}$  and  $j \neq i$ , respectively. The former corresponds to the inner bound approach that the encoders are assumed to operate independently, i.e., the joint conditional inner-bound pmf must factor. In contrast, the outer bound approach assumes that encoders can operate dependently on each other, which is a weaker condition on the outer-bound pmf.

The presented inner and outer bound in Theorem 1 and Theorem 2 coincide with known rate regions under certain conditions:

- If  $D_i \rightarrow 0, i = 1, \dots, N$ ,  $\mathcal{R}_a(D_{\mathcal{L}})$  and  $\mathcal{R}_c(D_{\mathcal{L}})$  coincide and reduce to the Slepian-Wolf rate region in [1].
- If only two sources are available  $\mathcal{R}_a(D_{\mathcal{L}})$  and  $\mathcal{R}_c(D_{\mathcal{L}})$  coincide with the inner and the outer bound presented in [8].
- If  $R_1 \geq H(X_1)$ ,  $\mathcal{R}_a(D_{\mathcal{L}})$  and  $\mathcal{R}_c(D_{\mathcal{L}})$  coincide with the inner and the outer bound presented in [13].

### D. Independently Degraded Helpers

In this section we present an inner and an outer bound for the rate-distortion region under the CI condition. The results herein are along the line of the reasoning in [13, Section V]. Eventually, we end up with a similar formulation, except that in our case the auxiliary sources are conditionally independent given the primary source, which is encoded with rate  $R_1$ , whereas in [13, Section V] the sources are conditionally independent given an side information, which is available at the receiver and not encoded. Due to this difference, our rate region  $\mathcal{R}_a(D_{\mathcal{L}})$  and  $\mathcal{R}_c(D_{\mathcal{L}})$  do not coincide in general and are only tight for specific constraints, as opposed to [13, Section V], where the inner and the outer bound are tight in general.

The CI assumption allows us to simplify the inner and the outer bound on the rate-distortion region, as follows.

*Corollary 3:* If  $X_2, X_3, \dots, X_N$  are conditionally independent given  $X_1$ , i.e.,  $X_{\mathcal{N}} \sim (4)$ , then

$$\mathcal{R}_a(D_{\mathcal{L}}) \subseteq \mathcal{R}(D_{\mathcal{L}}),$$

where  $\mathcal{R}_a(D_{\mathcal{L}})$  is the set of all rate  $N$ -tuples  $R_{\mathcal{N}}$  such that there exists a  $N - 1$ -tuple  $U_{\mathcal{L}}$  of discrete random variables with

$$p(u_{\mathcal{L}}, x_{\mathcal{N}}) = p(x_1) \prod_{i=2}^N p(u_i|x_i) \cdot p(x_i|x_1) \quad (13)$$

for which the following conditions are satisfied:

$$\begin{cases} \sum_{i \in \mathcal{S}} R_i \geq H(X_1|U_{\mathcal{S}^c}) + \sum_{i \in \mathcal{S} \setminus \{1\}} I(X_i; U_i|X_1) & \text{if } 1 \in \mathcal{S}, \\ R_i \geq I(X_i; U_i|X_1) & \forall i \in \mathcal{L}, \end{cases} \quad (14a)$$

$\forall \mathcal{S} \subseteq \mathcal{N}$ ; and for which there exist functions  $g_i(\cdot), \forall i \in \mathcal{N}$ , such that

$$\mathbb{E}[d_1(X_1, g_1(X_1, U_{\mathcal{L}}))] \leq \epsilon \quad (15)$$

and  $\forall i \in \mathcal{L}$

$$\mathbb{E}[d_i(X_i, g_i(X_1, U_i))] \leq D_i + \epsilon. \quad (16)$$

*Proof:* See Appendix C. ■

Note that  $g_i(\cdot), \forall i \in \mathcal{L}$  is only a function of  $X_1$  and  $U_i$ , in contrast to  $g_i(\cdot)$  in (8) where the function is dependent on  $X_1$  and all  $U_{\mathcal{L}}$ . The function  $g_1(\cdot)$  is still a function of  $X_1$  and  $U_{\mathcal{L}}$ .

*Corollary 4:* If  $X_2, X_3, \dots, X_N$  are conditionally independent given  $X_1$ , i.e.,  $X_{\mathcal{N}} \sim (4)$ , then  $\mathcal{R}'_c(D_{\mathcal{L}}) \supseteq \mathcal{R}_c(D_{\mathcal{L}})$ , and hence  $\mathcal{R}'_c(D_{\mathcal{L}}) \supseteq \mathcal{R}(D_{\mathcal{L}})$ , where  $\mathcal{R}'_c(D_{\mathcal{L}})$  is the set of all rate  $N$ -tuples  $R_{\mathcal{N}}$  such that there exists a  $N - 1$ -tuple  $U_{\mathcal{L}}$  of discrete random variables whose joint pmf  $p(u_{\mathcal{L}}, x_{\mathcal{N}})$  must satisfy the marginal constraints

$$\sum_{u_j, j \in \mathcal{L} \setminus \{i\}} p(u_{\mathcal{L}}, x_{\mathcal{N}}) = p(x_{\mathcal{N}})p(u_i|x_i), \quad (17)$$

$\forall i \in \mathcal{L}$ , for which the following conditions are satisfied:

$$\begin{cases} \sum_{i \in \mathcal{S}} R_i \geq H(X_1|U_{\mathcal{S}^c}) + \sum_{i \in \mathcal{S} \setminus \{1\}} I(X_i; U_i|X_1) & \text{if } 1 \in \mathcal{S}, \\ R_i \geq I(X_i; U_i|X_1) & \forall i \in \mathcal{L}, \end{cases} \quad (18a)$$

$$\quad (18b)$$

$\forall \mathcal{S} \subseteq \mathcal{N}$ ; and for which there exist functions  $g_i(\cdot), \forall i \in \mathcal{N}$ , such that

$$\mathbb{E}[d_1(X_1, g_1(X_1, U_{\mathcal{L}}))] \leq \epsilon \quad (19)$$

and  $\forall i \in \mathcal{L}$

$$\mathbb{E}[d_i(X_i, g_i(X_1, U_i))] \leq D_i + \epsilon. \quad (20)$$

*Proof:* See Appendix D. ■

Note that  $g_i(\cdot), \forall i \in \mathcal{L}$  is only a function of  $X_1$  and  $U_i$  in contrast to  $g_i(\cdot)$  in (12) where the function is dependent on  $X_1$  and all  $U_{\mathcal{L}}$ . The function  $g_1(\cdot)$  is still a function of  $X_1$  and  $U_{\mathcal{L}}$ .

The inner bound  $\mathcal{R}_a(D_{\mathcal{L}})$  in (14a), (14b) and the outer bound  $\mathcal{R}_c(D_{\mathcal{L}})$  in (18a), (18b) look similar, *except* that the underlying auxiliary random variables  $U_{\mathcal{L}}$  in Corollary 4 have more degrees of freedom, as pointed out in Section II-C. For (18b), each mutual information  $I(X_i; U_i|X_1)$  depends only on the marginal distribution of  $(X_i, U_i, X_1)$ ,  $\forall i \in \mathcal{L}$ . Hence the additional degrees of freedom cannot lower the mutual information. Therefore, the constraints on the rates in (14b) and (18b) are equivalent. However, this statement is not true for the constraints in (14a) and (18a). In Corollary 4, the additional degrees of freedom for the auxiliary random variables  $U_{\mathcal{L}}$  can result in a smaller conditional entropy  $H(X_1|U_{\mathcal{S}^c})$  in (18a) in comparison to the conditional entropy  $H(X_1|U_{\mathcal{S}^c})$  in (14a) of Corollary 3. In conclusion, the rate regions in Corollary 3 and Corollary 4 do not coincide in general, but they do coincide for (14b) and (18b).

*Remark:* If  $R_1 \geq H(X_1)$ , all sum rate bounds in (14a) and (18a) are implicitly satisfied by the lower single rate bounds of (14b) and (18b), respectively, i.e., all single rates  $R_i \geq I(X_i; U_i|X_1)$ . Hence, in this case  $\mathcal{R}_a(D_{\mathcal{L}}) = \mathcal{R}_c(D_{\mathcal{L}})$ , as presented in [13, Theorem 6].



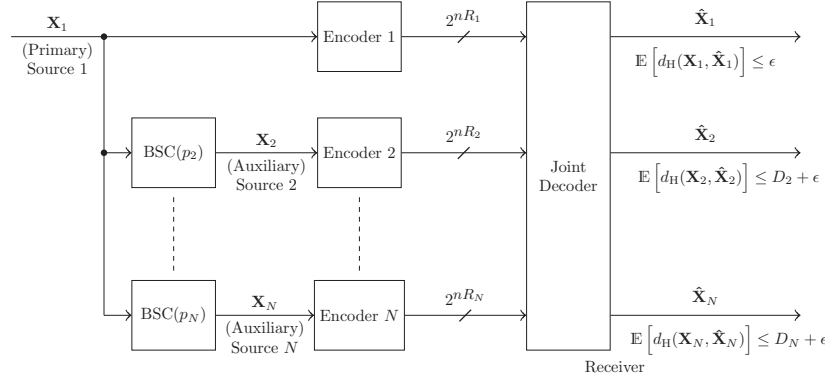


Fig. 2: The strong many-help-one problem with binary sources, CI condition, and Hamming distortion measure.

### E. Binary Sources and Hamming Distortion

We now consider a special case of the strong many-help-one problem, with binary sources, CI condition, and Hamming distortion measure illustrated in Fig. 2. Suppose that the random variable  $X_1$  takes values from a binary set  $\mathcal{X} = \{0, 1\}$  with uniform probabilities, i.e.,  $\Pr[X_1 = 0] = \Pr[X_1 = 1] = 0.5$ . Assume that  $\mathbf{X}_1$  is the input and  $\mathbf{X}_i$  is the output of a memoryless BSC with cross-over probability  $p_i \leq 0.5$ ,  $\forall i \in \mathcal{L}$ . All BSCs are independent, i.e.,  $p(\mathbf{x}_2, \dots, \mathbf{x}_N | \mathbf{x}_1) = \prod_{i=2}^N p(\mathbf{x}_i | \mathbf{x}_1)$ , and the errors occurring in  $\mathbf{X}_i$ ,  $\forall i \in \mathcal{L}$  are i.i.d., i.e.,  $p(\mathbf{x}_i | \mathbf{x}_1) = \prod_{k=1}^n p(x_{i,k} | x_{1,k})$ . Thus, the auxiliary sources  $X_i$ ,  $\forall i \in \mathcal{L}$ , are conditionally independent given the primary source. The distortion constraints are given by (1) and (2), where the distortion is measured by the Hamming distance, i.e.,

$$d_H(x_{i,k}, \hat{x}_{i,k}) = \begin{cases} 1, & \text{if } x_{i,k} \neq \hat{x}_{i,k}, \\ 0, & \text{if } x_{i,k} = \hat{x}_{i,k}. \end{cases} \quad (21)$$

In this section we provide an inner and an outer bound for the associated rate-distortion region. Single-letter expressions are found for the mutual information  $I(X_i; U_i | X_1)$ ,  $\forall i \in \mathcal{L}$  and for the conditional entropy  $H(X_1 | U_{S^c})$  in Corollary 3 and Corollary 4, in terms of the cross-over probability  $p_i$  and the distortion constraint  $D_i$ ,  $\forall i \in \mathcal{L}$ . To derive  $\min I(X_i; U_i | X_1)$  where the minimum is taken over all  $p(u_i | x_i)$  such that  $\mathbb{E}[d_i(X_i, g_i(X_1, U_i))] \leq D_i + \epsilon$ , we apply known results from [5] and [18, Theorem 10.3.1]. Depending on the rate  $R_i$ , we need to distinguish between two cases for the sake of completeness, as shall become apparent soon. In case one, referred to as joint decoding, each auxiliary source  $X_i$  is decoded *with* the help of the primary source  $X_1$ , i.e.,  $I(X_i; U_i | X_1)$  is at play; in case two, referred to as independent decoding, the auxiliary sources are decoded *without* the help of the primary source, i.e.,  $I(X_i; U_i)$  is at play. In Lemma 5 we derive a single-letter expression for  $\min H(X_1 | U_{S^c})$ , where the minimum is taken over all  $p(u_{S^c} | x_1)$  such that  $\mathbb{E}[d_H(X_1, g_1(X_1, U_{S^c}))] \leq \epsilon$ , based on [19].

#### 1) Mutual Information $I(X_i; U_i | X_1)$ :

a) *Joint Decoding*: If the auxiliary source  $X_i$  is decoded *with* the help of the primary source  $X_1$ , we can consider this as the Wyner-Ziv problem with side information at the decoder [5]. Here the primary source acts as the side information and is available at the receiver side. In [5, Section II] the binary case with uniform probabilities and Hamming distortion measure was investigated. From [5, Section II], the mutual information  $I(X_i; U_i | X_1)$  is given as

$$R'_i(D_i) = I(X_i; U_i | X_1) = \begin{cases} f(D_i) & \text{for } 0 \leq D_i \leq D_c, \\ (p_i - D_i)f'(D_c) & \text{for } D_c < D_i \leq p_i, \end{cases} \quad (22)$$

with  $f(D_i) = h(p_i * D_i) - h(D_i)$ ,  $f'(x) = \frac{\partial f(x)}{\partial x}$ , and  $D_c$  is the solution to the equation  $f(D_c)/(p_i - D_c) = f'(D_c)$ . In Fig. 3 the rate-distortion function of the Wyner-Ziv problem with side information at the decoder is presented. Additionally, an upper bound  $h(p_i * D_i) - h(D_i)$  and a lower bound  $h(p_i) - h(D_i)$  on  $R'_i(D_i)$  are shown. From now on, we refer to  $R'_i(D_i)$  as the joint decoding rate function. Please note that  $D_i$  is upper bounded by  $p_i$ .

So far we considered the reconstruction of the auxiliary source  $X_i$  under distortion constraint  $D_i$ . However, if encoder  $i$  provides a rate larger than  $1 - h(D_i)$  a distortion constraint  $d_i$  smaller than  $D_i$  can be achieved by decoding the auxiliary source  $X_i$  without any side information. In what follows we discuss this case.

b) *Independent Decoding*: We consider the case in which the auxiliary source  $X_i$  is decoded *without* the help of the primary source. We define  $U_i(d_i)$ ,  $0 \leq d_i \leq 0.5$ , to be a binary random variable obtained by connection to  $X_i$  via a BSC with

cross-over probability  $d_i$ . The auxiliary source  $X_i$  is decoded without any side information at the receiver, i.e.,  $d_i$  is not upper bounded by  $p_i$  as in (22). From [18, Theorem 10.3.1], the mutual information  $I(X_i; U_i)$  is given as

$$R_i(d_i) = I(X_i; U_i) = 1 - h(d_i) \text{ for } 0 \leq d_i \leq 0.5. \quad (23)$$

From now on, we refer to this function as independent decoding rate function, shown in Fig. 3. The mutual information  $R'_i(D_i) = I(X_i; U_i | X_1)$  in (22) is always upper bounded by the mutual information  $R_i(d_i) = I(X_i; U_i)$  in (23) for equal distortion constraints.

In summary, we found single-letter expressions for the mutual information in (14a), (14b), (18a), and (18b), in terms of the cross-over probability and the distortion constraint. Depending on the rate of encoder  $i$ , a distinction between joint and independent decoding is necessary. We now derive a single-letter expression for the conditional entropy  $H(X_i | U_{\mathcal{S}^c})$  in (14a) and (18a).

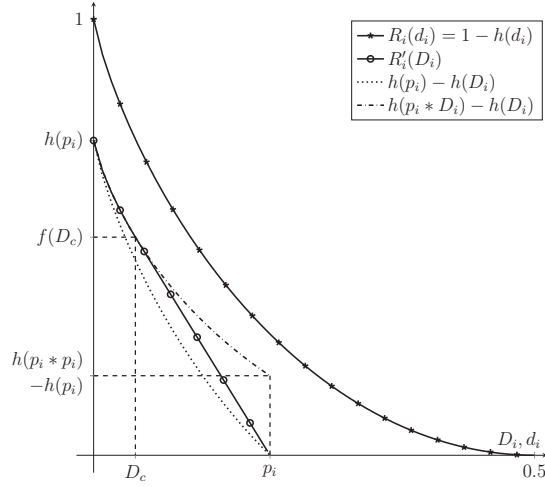


Fig. 3: Joint decoding rate function based on the Wyner-Ziv theorem and independent decoding rate function based on the rate-distortion theory.

2) *Conditional Entropy  $H(X_1 | U_{\mathcal{S}^c})$* : From the auxiliary sources  $X_i, \forall i \in \mathcal{S}^c$ ,  $U_{\mathcal{S}^c}$  is made available at the receiver, and we aim to reconstruct  $X_1$  perfectly. From (16) and (20) we know that the function  $g_i(\cdot), \forall i \in \mathcal{S}^c$  is only dependent on  $X_1$  and  $U_i$ . In other words, with the CI condition optimal performance is achieved if all auxiliary source encoders and decoders operate independently of each other. As  $H(X_1 | U_{\mathcal{S}^c})$  expresses the remaining uncertainty of  $X_1$  when  $U_{\mathcal{S}^c}$  is known, we can evaluate  $H(X_1 | U_{\mathcal{S}^c})$  by assuming that  $U_i, \forall i \in \mathcal{S}^c$  is decoded independently of  $X_1$  at the receiver. The problem of perfect reconstruction of source  $X_1$  with given side information  $U_{\mathcal{S}^c}$  corresponds to the source coding problem with side information, for which the rate region is given in [19]. In [19, Section II] it is shown that  $\min H(X_1 | U_{\mathcal{S}^c})$  is achieved if the primary source encoder operates independently from the encoder processing the side information. In our case the side information is  $U_{\mathcal{S}^c}$  processed by encoders  $i$  for  $i \in \mathcal{S}^c$ . Thus, if the auxiliary random variables  $U_{\mathcal{S}^c}$  satisfy

$$p(x_1, x_{\mathcal{S}^c}, u_{\mathcal{S}^c}) = p(x_1, x_{\mathcal{S}^c}) \prod_{i \in \mathcal{S}^c} p(u_i | x_i), \quad (24)$$

we achieve  $\min H(X_1 | U_{\mathcal{S}^c})$ , where the minimum is taken over all  $p(u_{\mathcal{S}^c} | x_1)$  such that  $\mathbb{E}[d_H(X_1, g_1(X_1, U_{\mathcal{S}^c}))] \leq \epsilon$ . In summary, we can make the following statements for the conditional entropy  $H(X_1 | U_{\mathcal{S}^c})$  considering binary sequences, CI condition and the Hamming distortion measure.

- 1)  $\min H(X_1 | U_{\mathcal{S}^c})$  is achieved if all encoders operate independently.
- 2) As a consequence of 1), the inner bound in Corollary 3 and the outer bound in Corollary 4 for  $\min H(X_1 | U_{\mathcal{S}^c})$  coincide.
- 3) Auxiliary source decoders operate independently.

We now derive the single-letter expression for the conditional entropy.

**Lemma 5:** Let  $X_{v_1}, X_{v_2}, \dots, X_{v_{|\mathcal{V}|}}$  be conditionally independent given  $X_1$  which takes values from the binary set  $\mathcal{X} = \{0, 1\}$  with uniform probabilities. The dependency between  $X_1$  and  $X_v$  is given by the cross-over probability  $p_v \in [0, 0.5]$  of a BSC, for  $v \in \mathcal{V}$  and arbitrary set  $\mathcal{V}$ . The binary auxiliary random variables  $U_{\mathcal{V}}$  are constructed according to the pmf in (24). Furthermore, sources  $X_v, v \in \mathcal{V}$ , are decoded *without* any side information at the receiver, i.e., the rate-distortion function for



each source  $X_v$ ,  $v \in \mathcal{V}$ , is given by (23). The dependency between  $X_v$  and  $U_v$  is given by the distortion constraint  $d_v \in [0, 0.5]$ , for  $v \in \mathcal{V}$ . Moreover, the rate function for the primary source is

$$R_1 = \min_{p(\hat{x}_1|x_1): \mathbb{E}[d_H(x_1, \hat{x}_1)] \leq \epsilon} I(X_1; \hat{X}_1) = \min_{p(u_{\mathcal{V}}|x_1): \mathbb{E}[d_H(x_1, g_1(x_1, u_{\mathcal{V}}))] \leq \epsilon} H(X_1|U_{\mathcal{V}}) = \phi(p_{\mathcal{V}}, d_{\mathcal{V}}), \quad (25)$$

where  $\phi(p_{\mathcal{V}}, d_{\mathcal{V}})$  is the single-letter expression of  $H(X_1|U_{\mathcal{V}})$  in terms of cross-over probabilities  $p_{\mathcal{V}}$  and distortion constraints  $d_{\mathcal{V}}$  given in (28).

*Proof:*  $\min H(X_1|U_{\mathcal{V}})$  is achieved for independent encoders as argued above, i.e.,  $X_1, U_{\mathcal{V}} \sim (24)$ . Thus, we can rewrite the conditional entropy in (25) as follows

$$\min H(X_1|U_{\mathcal{V}}) = H(X_1, U_{\mathcal{V}}) - H(U_{\mathcal{V}}) \quad (26)$$

$$\begin{aligned} &= H(X_1) + \sum_{\{v_1\} \in \mathcal{V}} H(U_{v_1}|X_1) - \sum_{\{v_1\} \in \mathcal{V}} H(U_{v_1}) + \sum_{\{v_1, v_2\} \in \mathcal{V}} I(U_{v_1}; U_{v_2}) \\ &\quad - \sum_{\{v_1, v_2, v_3\} \in \mathcal{V}} I(U_{v_1}; U_{v_2}; U_{v_3}) + \dots \pm \sum_{\{v_1, \dots, v_{|\mathcal{V}|}\} \in \mathcal{V}} I(U_{v_1}; \dots; U_{v_{|\mathcal{V}|}}). \end{aligned} \quad (27)$$

If  $\mathcal{V} \neq \{\emptyset\}$ , then

$$\begin{aligned} \min H(X_1|U_{\mathcal{V}}) &= \sum_{v_1 \in \mathcal{V}} h(p_{v_1} * d_{v_1}) - \sum_{\{v_1, v_2\} \in \mathcal{V}} h(p_{v_1} * d_{v_1} * p_{v_2} * d_{v_2}) \\ &\quad + \sum_{\{v_1, v_2, v_3\} \in \mathcal{V}} h(p_{v_1} * d_{v_1} * p_{v_2} * d_{v_2} * p_{v_3} * d_{v_3}) \\ &\quad - \dots \pm \sum_{\{v_1, \dots, v_{|\mathcal{V}|}\} \in \mathcal{V}} h(p_{v_1} * d_{v_1} * \dots * p_{v_{|\mathcal{V}|}} * d_{v_{|\mathcal{V}|}}) \end{aligned} \quad (28)$$

$$\triangleq \phi(p_{\mathcal{V}}, d_{\mathcal{V}}). \quad (29)$$

If  $\mathcal{V} = \{\emptyset\}$ , then

$$\min H(X_1|\emptyset) = H(X_1) = 1 \triangleq \phi(\emptyset, \emptyset). \quad (30)$$

The steps are justified as follows: (26) is the chain rule of entropy; for (27) we have used the chain rule of mutual information and multivariate mutual information [20] and the fact that  $U_{v_1}, U_{v_2}, \dots, U_{v_{|\mathcal{V}|}}$  are conditionally independent given  $X_1$ ; and (28) follows from the Markov chain  $U_v \rightarrow X_v \rightarrow X_1$ , with use of  $H(U_v|X_1) = h(p_v * d_v)$  as proven in [19]. From the Markov chain  $U_{v_i} \rightarrow X_{v_i} \rightarrow X_1 \rightarrow X_{v_j} \rightarrow U_{v_j}$  (following from (24)) we have the conditional entropy  $H(U_{v_i}|U_{v_j}) = h(p_{v_i} * d_{v_i} * p_{v_j} * d_{v_j})$  for  $v_i \in \mathcal{V}$ ,  $v_j \in \mathcal{V}$  and  $v_i \neq v_j$ . From the assumption on the pmf in (24) we also have the multivariate mutual information given as  $I(U_{v_1}; \dots; U_{v_{|\mathcal{V}|}}) = H(X_1) - h(p_{v_1} * d_{v_1} * \dots * p_{v_{|\mathcal{V}|}} * d_{v_{|\mathcal{V}|}})$ . After some algebraic manipulations omitted here we finally arrive at (28). This completes the proof. ■

In what follows we have to carefully distinguish between jointly decoding the auxiliary sources  $X_i$  *with* “side information”  $X_1$  and independently decoding the auxiliary sources  $X_i$  *without* “side information”  $X_1$ ,  $\forall i \in \mathcal{L}$ . Eventually, we have to determine which decoding strategy minimizes and maximizes the inner and the outer bound — including both the conditional entropy and the mutual information — in (14a), (14b), (18a), and (18b).

3) *Outer Bound:* We now derive a single-letter expression for Corollary 4 with binary source, CI condition, and Hamming distortion measure.

*Theorem 6:* If  $X_1$  takes values from a binary set with uniform probabilities,  $X_{\mathcal{N}} \sim p(x_{\mathcal{N}}) = p(x_1) \prod_{i=2}^N p(x_i|x_1)$  and the binary random variables  $U_{\mathcal{L}}$  are distributed with pmf (17), then it holds that

$$\mathcal{R}_c(D_{\mathcal{L}}) \supseteq \mathcal{R}(D_{\mathcal{L}}),$$

where  $\mathcal{R}_c(D_{\mathcal{L}})$  is the set of all rate  $N$ -tuples  $R_{\mathcal{N}}$  for which the following conditions are satisfied:

$$\begin{cases} \sum_{i \in \mathcal{S}} R_i \geq \phi(p_{\mathcal{S}^c}, d_{\mathcal{S}^c}) + \sum_{i \in \mathcal{S} \setminus \{1\}} R'_i(D_i) & \text{if } 1 \in \mathcal{S}, \\ R_i \geq R_i^*(d_i, \rho_i) & \forall i \in \mathcal{L}, \end{cases} \quad (31a)$$

$$\quad (31b)$$

where

$$R_i^*(d_i, \rho_i) = \begin{cases} 1 - h(d_i) & \text{for } 0 \leq d_i < D_i, \\ \arg \max_{R_i = 1 - h(d_i)} \mathcal{R}_c(D_{\mathcal{L}}) & \text{for } 0 \leq \rho_i \leq 1, D_i \leq d_i \leq 0.5, \\ R_i = \rho_i \cdot (1 - h(D_i)) + (1 - \rho_i) \cdot R'_i(D_i) & \end{cases} \quad (32a)$$

$$\quad (32b)$$

$\forall i \in \mathcal{L}$ .

*Proof:* Assume that a  $D_i$ -admissible code exists. We can make the following statements.

- The mutual information  $I(X_i; U_i | X_1)$  in (18a), Corollary 4, is at least  $R'_i(D_i)$ , as shown in Section II-E1, Paragraph a) i.e., the auxiliary source is decoded with the primary source.
- The conditional entropy  $H(X_1 | U_{\mathcal{S}^c})$  in (18a), Corollary 4, is at least  $\phi(p_{\mathcal{S}^c}, d_{\mathcal{S}^c})$  as proved in Lemma 5, (25).
- For the mutual information  $I(X_i; U_i | X_1)$  in (18b), Corollary 4, two decoding strategies can be considered. Other than the mutual information in (18a)<sup>1</sup>, the mutual information in (18b) can either be determined by joint decoding, i.e.,  $I(X_i; U_i | X_1) = R'_i(D_i)$  (Section II-E1, Paragraph a)), or independent decoding, i.e.,  $I(X_i; U_i) = 1 - h(d_i)$  (Section II-E1, Paragraph b)). Ultimately it is crucial which decoding strategy maximizes  $\mathcal{R}_c(D_{\mathcal{L}})$ . Thus the mutual information in (18b) is at least  $R_i^*(d_i, \rho_i)$ .

This completes the proof. ■

Note that in (31a) the underlying distribution of  $\mathbf{U}_{\mathcal{S}^c}$  is (24) as discussed in Lemma 5. The function  $R_i^*(d_i, \rho_i)$  is split into two regions:

- (32a): If the rate is  $R_i^* > 1 - h(D_i)$ , a distortion constraint  $d_i$  that is inferior to  $D_i$  can be achieved by independent decoding, i.e.,  $I(X_i; U_i)$ , such that  $R_i^* = 1 - h(d_i)$  for  $0 \leq d_i < D_i$ . In other words, independent decoding maximizes the outer bound  $\mathcal{R}_c(D_{\mathcal{L}})$ .
- (32b): If the rate is  $1 - h(D_i) \geq R_i^* \geq R'_i(D_i)$  joint decoding, i.e.,  $I(X_i; U_i | X_1)$ , or independent decoding, i.e.,  $I(X_i; U_i)$ , maximizes the outer bound  $\mathcal{R}_c(D_{\mathcal{L}})$ . If joint decoding maximizes  $\mathcal{R}_c(D_{\mathcal{L}})$ , then  $R_i^*(\cdot, \rho_i) = \rho_i \cdot (1 - h(D_i)) + (1 - \rho_i) \cdot R'_i(D_i)$  and if independent decoding maximizes  $\mathcal{R}_c(D_{\mathcal{L}})$ , then  $R_i^*(d_i, \cdot) = 1 - h(d_i)$ .

Note that  $\rho_i$  and  $d_i$  are continuous parameters that directly affect the outer bound  $\mathcal{R}_c(D_{\mathcal{L}})$ . In Section II-E5 we discuss and illustrate results of the outer bound for two and three sources.

4) *Inner Bound:* We now derive a single-letter expression for Corollary 3 with binary source, CI condition, and Hamming distortion measure.

*Theorem 7:* If  $X_1$  takes values from a binary set with uniform probabilities,  $X_{\mathcal{N}} \sim p(x_{\mathcal{N}}) = p(x_1) \prod_{i=2}^N p(x_i | x_1)$  and the binary random variables  $U_{\mathcal{L}}$  are distributed with pmf (13), then it holds that

$$\mathcal{R}_a(D_{\mathcal{L}}) \subseteq \mathcal{R}(D_{\mathcal{L}}),$$

where  $\mathcal{R}_a(D_{\mathcal{L}})$  is the set of all rate  $N$ -tuples  $R_{\mathcal{N}}$  for which the following conditions are satisfied:

$$R_1 \geq \begin{cases} \phi(p_{\mathcal{L}}, d_{\mathcal{L}}) & \text{for } 0 \leq d_i < D_i, \forall i \in \mathcal{L}, \\ \arg \text{conv}_{R_1} \{ \xi(p_{\mathcal{L}}, d_{\mathcal{Q}^c}, D_{\mathcal{Q}}, \rho_{\mathcal{Q}}) \} & \text{for } 0 \leq d_i < D_i, \forall i \in \mathcal{Q}^c, \rho_i = \{0, 1\}, \forall i \in \mathcal{Q}, \end{cases} \quad (33a)$$

$$R_i \geq \begin{cases} 1 - h(d_i) & \text{for } 0 \leq d_i < D_i, \forall i \in \mathcal{Q}^c, \\ \rho_i \cdot (1 - h(D_i)) + (1 - \rho_i) \cdot R'_i(D_i) & \text{for } 0 \leq \rho_i \leq 1, \forall i \in \mathcal{Q}, \end{cases} \quad (34a)$$

$$(34b)$$

$\forall \mathcal{Q}^c \subset \mathcal{L}$ , with  $\mathcal{Q}$  being the complement of  $\mathcal{Q}^c$ .  $\{\xi(\cdot)\} = \{R_1, R_{\mathcal{Q}^c}, R_{\mathcal{Q}}\}$  represents all known achievable  $N$ -rate tuples, given in (37). Furthermore,  $\text{conv}$  is a *Simplex*, introduced in [21, Section 2.2.4], defined as

$$C = \text{conv}\{v_1, \dots, v_k\} = \left\{ \theta_1 v_1 + \dots + \theta_k v_k \mid \theta_i \geq 0, \sum_i \theta_i = 1 \right\}. \quad (35)$$

*Proof:* We know achievable rate  $N$ -tuples  $R_{\mathcal{N}}$  in Corollary 3 for particular rate  $N - 1$ -tuples  $R_{\mathcal{L}}$ . The mutual information in (14b) is either

- $R_i = I(X_i; U_i) = 1 - h(d_i)$ , for  $0 \leq d_i \leq D_i, \forall i \in \mathcal{Q}^c$ , i.e., (14b) = (34b) with  $\rho_i = 1$  (independent decoding); or
- $R_i = I(X_i; U_i | X_1) = R'_i(D_i), \forall i \in \mathcal{Q}$ , i.e., (14b) = (34a) (joint decoding).

The achievable rate  $R_1$  can be determinant with the corresponding achievable rate sum bound in (14a). The underlying assumption for the sum rate bound is that the primary source is jointly decoded with side information  $U_{\mathcal{Q}^c}$  and the auxiliary sources  $i, \forall i \in \mathcal{Q}$  are jointly decoded with the primary source  $X_1$ . The achievable rate sum bound in (14a) is

$$\sum_{i \in \{1\} \cap \mathcal{Q}} R_i = H(X_1 | U_{\mathcal{S}^c}) + \sum_{i \in \mathcal{Q}} I(X_i; U_i | X_1) = \phi(p_{\mathcal{Q}^c}, d_{\mathcal{Q}^c}) + \sum_{i \in \mathcal{Q}} R'_i(D_i), \quad (36)$$

<sup>1</sup>The mutual information in (18a) is considered in a sum rate bound, this assumes in principle a joint decoding strategy.

where the conditional entropy is given in Lemma 5. Thus,  $R_1 = \phi(p_{Q^c}, d_{Q^c})$  is known to be achievable. Therefore, the following rate  $N$ -tuples are known to be achievable

$$\{\xi(\cdot)\} = \{R_1, R_{Q^c}, R_Q\} = \left\{ \phi(p_{Q^c}, d_{Q^c}), \underbrace{1 - h(d_i) | 0 \leq d_i \leq D_i, \forall i \in Q^c}_{\text{independently decoded auxiliary sources, i.e., } I(X_i; U_i)}, \underbrace{R'_i(D_i) | \forall i \in Q}_{\text{jointly decoded auxiliary sources, i.e., } I(X_i; U_i | X_1)} \right\} \quad (37)$$

$\forall Q^c \subseteq \mathcal{L}$ . The *Simplex* forms a convex hull from all known achievable rate  $N$ -tuples  $\{\xi(\cdot)\}$ . All rate  $N$ -tuples inside the convex hull are achievable (this can be reasoned by the time-sharing argument). Then, we can determine the minimum rate  $R_1$  for a given rate  $N - 1$ -tuple  $R_{\mathcal{L}}$ . This completes the proof. ■

Note that in (33a) all auxiliary sources are individually decoded and thus the constraint for rate  $R_1$  is equivalent to Lemma 5. To illustrate this important fact we write this case separately although it could also be included in (33b). The auxiliary encoder rate range for joint decoding and individual decoding is determined in (34a) and (34b), respectively.<sup>2</sup> Section II-E5 will discuss and illustrate the inner bound for two and three sources.

5) *Discussion:* In this section, we illustrate and discuss the results from Theorem 6 and Theorem 7 for two and three sources.

a) *Two Sources:* From (31a) and (31b) the outer bound  $\mathcal{R}_c(D_2)$  for binary sources, CI condition, and Hamming distortion measure is given by

$$R_1 \geq H(X_1 | U_2) = \phi(p_2, d_2), \quad (38a)$$

$$R_2 \geq I(X_2, U_2 | X_1) = R_2^*(d_2, \rho_2), \quad (38b)$$

$$R_1 + R_2 \geq H(X_1) + I(X_2, U_2 | X_1) = 1 + R_2'(D_2). \quad (38c)$$

For an easier understanding we consider two other outer bounds, as follows. First, we determine the rate region  $\mathcal{R}_c^1$  that gives an outer bound for reconstruction of the auxiliary source with distortion  $D_2$ , i.e., ignoring rate constraint in (38a):

$$\mathcal{R}_c^1 = \{(R_1, R_2) : R_1 + R_2 \geq 1 + R_2'(D_2), R_2 \geq R_2'(D_2)\}. \quad (39)$$

Second, we determine a rate region  $\mathcal{R}_c^2$  that gives an outer bound for perfect reconstruction of the primary source, i.e., ignoring rate constraint in (38c):

$$\mathcal{R}_c^2 = \{(R_1, R_2) : R_1 \geq h(p_2 * d_2), R_2 \geq 1 - h(d_2) \text{ for } 0 \leq d_2 \leq 0.5\}. \quad (40)$$

The target outer bound  $\mathcal{R}_c(D_2)$  is the intersection of  $\mathcal{R}_c^1$  and  $\mathcal{R}_c^2$ .

From (33a) - (34b) the inner bound  $\mathcal{R}_a(D_2)$  for binary sources, CI condition, and Hamming distortion measure is given by

$$R_1 \geq \begin{cases} \phi(p_2, d_2) & \text{for } 0 \leq d_2 \leq D_2, \\ \arg \max_{R_1} \text{conv} \{\xi(p_2, \emptyset, D_2, \rho_2)\} & \text{for } \rho_2 = \{0, 1\}. \end{cases} \quad (41a)$$

$$R_2 \geq \begin{cases} 1 - h(d_2) & \text{for } 0 \leq d_2 \leq D_2, \\ \rho_2 \cdot (1 - h(D_2)) + (1 - \rho_2) \cdot R_2'(D_2) & \text{for } 0 \leq \rho_2 \leq 1. \end{cases} \quad (42a)$$

$$R_2 \geq \begin{cases} 1 - h(d_2) & \text{for } 0 \leq d_2 \leq D_2, \\ \rho_2 \cdot (1 - h(D_2)) + (1 - \rho_2) \cdot R_2'(D_2) & \text{for } 0 \leq \rho_2 \leq 1. \end{cases} \quad (42b)$$

The following two rate 2-tuples are known to be achievable: (i) if  $R_2 = R_2'(D_2)$  then we know from (37) that  $R_1 = \phi(\emptyset, \emptyset) = 1$  is achievable (i.e.,  $Q^c = \{\emptyset\}$ ) and, similarly, (ii) if  $R_2 = 1 - h(d_2)$ , for  $0 \leq d_2 \leq D_2$ , we know that  $R_1 = \phi(p_2, d_2) = h(p_2 * d_2)$  is achievable (i.e.,  $Q^c = \{2\}$ ).

In the following we consider two cases for different ranges of  $D_2$ .

*Case 1:* Fig. 4a illustrates the inner and the outer bound if  $D_2 \leq D_{2c}$ <sup>3</sup>:

- Inner bound: In Fig. 4a point  $A_1$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{1, h(p_2 * D_2) - h(D_2)\}$  and point  $B_1$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{h(p_2 * D_2), 1 - h(D_2)\}$ . Point  $A_1$  and  $B_1$  can be connected by the time-sharing argument. This is the inner bound  $\mathcal{R}_a^1$  for which it holds that the primary source is perfectly reconstructed and the auxiliary source is reconstructed with distortion constraint  $D_2$ . In this case, the auxiliary source is jointly decoded<sup>4</sup>. In Fig. 4c the rate  $R_2$  is illustrated. We see that point  $A_1$  is on the joint decoding rate function and point  $B_1$  is on the independent

<sup>2</sup>Note the difference between  $\mathcal{S}^c$  in Theorem 6 and  $Q^c$ . In Theorem 6, individually decoded sources are determined by the maximum argument in (32b). This can differ from Theorem 7, since  $d_i$  is not upper-bounded by  $D_i$  for independent decoding.

<sup>3</sup> $D_c$  results from the case distinction in (22).

<sup>4</sup>All auxiliary source rates between  $A_1$  and  $B_2$  are achieved by time sharing of independent decoding and joint decoding. Since reconstruction of the auxiliary source is not exclusively achieved by independent decoding we refer to this as joint decoding.

decoding rate function. An additional inner bound  $\mathcal{R}_a^2$  exists for which it holds that the primary source is perfectly reconstructed and the auxiliary source is reconstructed with distortion constraint  $d_2$  for  $0 \leq d_2 \leq D_2$ . In this case, the auxiliary source is independently decoded, thus the achievable rate 2-tuple is  $\{R_1, R_2\} = \{h(p_2 * d_2), 1 - h(d_2)\}$ . Where the achievable rate  $R_1$  is given by Lemma 5. In Fig. 4a point  $C$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{h(p_2), 1\}$  for  $d_2 = 0$ . We can connect point  $B_1$  and  $C$  for  $D_2 \geq d_2 \geq 0$  accordingly. The corresponding rate  $R_2$  is illustrated in Fig. 4c, where point  $B_1$  and point  $C$  are on the independent decoding rate function.

- Outer bound: Fig. 4a illustrates the outer bounds  $\mathcal{R}_c^1$  and  $\mathcal{R}_c^2$ . For  $D_2 \leq D_{2c}$ ,  $\mathcal{R}_c^1$  in (39) is equivalent to

$$\mathcal{R}_c^1 = \{(R_1, R_2) : R_1 + R_2 \geq 1 + h(p_2 * D_2) - h(D_2), R_2 \geq h(p_2 * D_2) - h(D_2)\}.$$

$\mathcal{R}_c^2$  is given in (40).

- In conclusion, for the range  $R_2 \in [R'_2(D_2), 1 - h(D_2)]$  the inner and the outer bound are  $\mathcal{R}_c = \max[\mathcal{R}_c^1, \mathcal{R}_c^2] = \mathcal{R}_c^1$  and  $\mathcal{R}_a = \mathcal{R}_a^1$  with distortion  $D_2$ . In this rate range, the auxiliary source is jointly decoded. Inner bound  $\mathcal{R}_a^2$  does not achieve a distortion constraint  $d_2$  less or equal to  $D_2$  if  $R_2 < 1 - h(D_2)$ . For the range  $R_2 \in (1 - h(D_2), 1]$  the inner and the outer bound are  $\mathcal{R}_c = \max[\mathcal{R}_c^1, \mathcal{R}_c^2] = \mathcal{R}_c^2$  and  $\mathcal{R}_a = \mathcal{R}_a^2$  with distortion constraint  $d_2 \in [0, D_2]$ . In this rate range, the auxiliary source is independently decoded. The inner and the outer bound are tight for the whole range of  $R_2$ .

*Case 2:* Fig. 4b illustrates the inner and the outer bound if  $D_{2c} < D_2 \leq p_2$ .

- Inner bound: In Fig. 4b point  $A_2$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{1, h(p_2 * D_2) - h(D_2)\}$  and point  $B_2$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{h(p_2 * D_2), 1 - h(D_2)\}$ . Again, point  $A_2$  and  $B_2$  can be connected the by time-sharing argument. This is the inner bound  $\mathcal{R}_a^1$ . The inner bound  $\mathcal{R}_a^2$  is given by  $\{R_1, R_2\} = \{h(p_2 * d_2), 1 - h(d_2)\}$ , as discussed in *Case 1*. We can connect point  $B_2$  and  $C$  for  $D_2 \geq d_2 \geq 0$  accordingly.
- Outer bound: Fig. 4b illustrates the outer bounds  $\mathcal{R}_c^1$  and  $\mathcal{R}_c^2$ . For  $D_{2c} \geq D_2 \geq p_2$ ,  $\mathcal{R}_c^1$  in (39) is equivalent to

$$\begin{aligned} \mathcal{R}_c^1 = \{(R_1, R_2) : R_1 + R_2 \geq 1 + \rho_2 \cdot (1 - h(D_2)) + (1 - \rho_2) \cdot R'_2(D_2), \\ R_2 \geq \rho_2 \cdot (1 - h(D_2)) + (1 - \rho_2) \cdot R'_2(D_2)\}, \end{aligned}$$

where  $\rho_2$  is a scalar, such that the joint decoding rate function in Fig. 4c achieves distortion  $D_2$  at point  $A_2$ .  $\mathcal{R}_c^2$  is given in (40).

- In conclusion, for the range  $R_2 \in [R'_2(D_2), 1 - h(D_2)]$  the inner and the outer bound are  $\mathcal{R}_c = \max[\mathcal{R}_c^1, \mathcal{R}_c^2]$  and  $\mathcal{R}_a = \mathcal{R}_a^1$  with distortion  $D_2$ . Please note the difference to *Case 1*. The outer bound  $\mathcal{R}_c^2$  can be tighter than  $\mathcal{R}_c^1$ , but a mismatch between  $\mathcal{R}_c$  and  $\mathcal{R}_a$  remains, depicted in Fig. 4b as hatched area. For the inner bound, joint decoding is performed for range  $R_2 \in [R'_2(D_2), 1 - h(D_2)]$ , in contrast to the outer bound, where independent decoding for some range (dotted area) is optimal. For the range  $R_2 \in (1 - h(D_2), 1]$  the inner and the outer bound are  $\mathcal{R}_c = \mathcal{R}_c^2$  and  $\mathcal{R}_a = \mathcal{R}_a^2$  with distortion  $d_2 \in [0, D_1]$ . The inner and the outer bound for the range  $R_2 \in (1 - h(D_2), 1]$  are tight, and independent decoding is performed.

*Remark:* This result corresponds to [8, Section VI]. Berger and Yeung even showed an additional achievable rate pair  $\{R_2, R_1\} = \{\bar{\kappa}(1 - h(D_c)), \kappa + \bar{\kappa}h(p_2 * D_c)\}$  with  $\kappa = (D_2 - D_{2c})/(p_2 - D_{2c})$ , thereby achieving closer bounds. However, also these inner and outer bound do not coincide for  $D_{2c} < D_2 < p_2$ . In this study, we do not consider this rate pair to achieve a generalized solution for  $N$  sources.

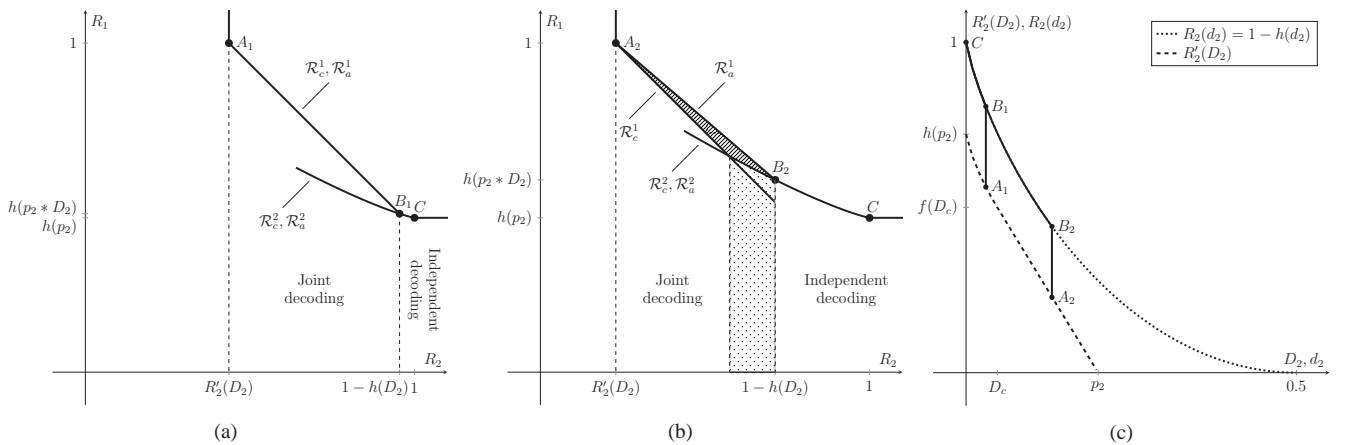


Fig. 4: The inner and the outer bound for binary sources, CI condition, and Hamming distortion measure with two sources: (a)  $0 \leq D_2 \leq D_{2c}$ , (b)  $D_{2c} < D_2 \leq p_2$ , and (c) rate function for second source under independent and joint decoding. *Remark:* both outer bounds,  $\mathcal{R}_c^1$  and  $\mathcal{R}_c^2$ , are drawn incompletely for the sake of clarity.

b) *Three Sources*: From (31a) and (31b) the outer bound  $\mathcal{R}_c(D_2, D_3)$  for binary sources, CI condition, and Hamming distortion measure is given by

$$R_1 \geq \phi(p_2, p_3, d_2, d_3), \quad (43a)$$

$$R_2 \geq R_2^*(d_2, \rho_3), \quad (43b)$$

$$R_3 \geq R_3^*(d_3, \rho_3), \quad (43c)$$

$$R_1 + R_2 \geq \phi(p_3, d_3) + R_2'(D_2), \quad (43d)$$

$$R_1 + R_3 \geq \phi(p_2, d_2) + R_3'(D_3), \quad (43e)$$

$$R_1 + R_2 + R_3 \geq 1 + R_2'(D_2) + R_3'(D_3). \quad (43f)$$

We like to point out that we can consider six outer bound<sup>5</sup>, the intersection of which is the target outer bound  $\mathcal{R}_c(D_2, D_3)$ . However, we don't specifically give these six outer bounds but rather discuss the outer bound in detail for different rate regions subsequently.

From (33a) - (34b) the inner bound  $\mathcal{R}_a(D_2, D_3)$  for binary sources, CI condition, and Hamming distortion measure is given by

$$R_1 \geq \begin{cases} \phi(p_2, p_3, d_2, d_3) & \text{for } 0 \leq d_i \leq D_i, i = \{2, 3\}, \\ \arg \text{conv}_{R_1} [\xi(p_2, p_3, d_2, D_3, \rho_3)] & \text{for } 0 \leq d_2 \leq D_2, \rho_3 = \{0, 1\}, \\ \arg \text{conv}_{R_1} [\xi(p_2, p_3, d_3, D_2, \rho_2)] & \text{for } 0 \leq d_3 \leq D_3, \rho_2 = \{0, 1\}, \\ \arg \text{conv}_{R_1} [\xi(p_2, p_3, D_2, D_3, \rho_2, \rho_3)] & \text{for } \rho_i = \{0, 1\}, i = \{2, 3\}, \end{cases} \quad (44a)$$

$$R_1 \geq \begin{cases} \arg \text{conv}_{R_1} [\xi(p_2, p_3, d_2, D_3, \rho_3)] & \text{for } 0 \leq d_2 \leq D_2, \rho_3 = \{0, 1\}, \\ \arg \text{conv}_{R_1} [\xi(p_2, p_3, d_3, D_2, \rho_2)] & \text{for } 0 \leq d_3 \leq D_3, \rho_2 = \{0, 1\}, \\ \arg \text{conv}_{R_1} [\xi(p_2, p_3, D_2, D_3, \rho_2, \rho_3)] & \text{for } \rho_i = \{0, 1\}, i = \{2, 3\}, \end{cases} \quad (44b)$$

$$R_1 \geq \begin{cases} \arg \text{conv}_{R_1} [\xi(p_2, p_3, d_3, D_2, \rho_2)] & \text{for } 0 \leq d_3 \leq D_3, \rho_2 = \{0, 1\}, \\ \arg \text{conv}_{R_1} [\xi(p_2, p_3, D_2, D_3, \rho_2, \rho_3)] & \text{for } \rho_i = \{0, 1\}, i = \{2, 3\}, \end{cases} \quad (44c)$$

$$R_1 \geq \begin{cases} \arg \text{conv}_{R_1} [\xi(p_2, p_3, D_2, D_3, \rho_2, \rho_3)] & \text{for } \rho_i = \{0, 1\}, i = \{2, 3\}, \end{cases} \quad (44d)$$

$$R_i \geq \begin{cases} 1 - h(d_i) & \text{for } 0 \leq d_i \leq D_i, \\ \rho_i \cdot (1 - h(D_i)) + (1 - \rho_i) \cdot R_i'(D_i) & \text{for } 0 \leq \rho_i \leq 1. \end{cases} \quad (45a)$$

$$R_i \geq \begin{cases} \rho_i \cdot (1 - h(D_i)) + (1 - \rho_i) \cdot R_i'(D_i) & \text{for } 0 \leq \rho_i \leq 1. \end{cases} \quad (45b)$$

for  $i = 2, 3$ . We have separated the rate constraints for rate  $R_1$  for easier understanding. Note the difference between the distortion constraints  $d_i$  and  $D_i$ . Distortion constraint  $d_i \in [0, D_i]$  is defined for a certain interval, whereas distortion constraint  $D_i$  is fixed.

- The rate constraint in (44a) corresponds to the known achievable rate 3-tuple  $\{\xi(\cdot)\} =$

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, d_2, d_3), 1 - h(d_2), 1 - h(d_3)\} \quad (46)$$

for which both auxiliary sources are decoded without the primary source, i.e.,  $R_i = 1 - h(d_i)$  for  $0 \leq d_i \leq D_i, i = \{2, 3\}$ .

- The rate constraint in (44b) forms a convex hull between the two known achievable rate 3-tuples  $\{\xi(\cdot)\} =$

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, d_2, D_3), 1 - h(d_2), 1 - h(D_3)\}, \quad (47a)$$

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, D_2, D_3), 1 - h(D_2), 1 - h(D_3)\}, \quad (47b)$$

$0 \leq d_2 \leq D_2$  for which the auxiliary source two is decoded without the primary source, i.e.,  $R_2 = 1 - h(d_2)$ ,  $0 \leq d_2 \leq D_2$  and the auxiliary source three is decoded with the primary source, i.e.,  $R_3 = \rho_3 \cdot (1 - h(D_3)) + (1 - \rho_3) \cdot R_3'(D_3)$ .

- The rate constraint in (44c) is symmetric to (44b) with interchanged indexes 2 and 3.
- The rate constraint in (44d) forms a convex hull between four known achievable rate 3-tuples  $\{\xi(\cdot)\} =$

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, D_2, D_3), 1 - h(D_2), 1 - h(D_3)\}, \quad (48a)$$

$$\{R_1, R_2, R_3\} = \{\phi(p_3, p_2, D_3, D_2), 1 - h(D_3), 1 - h(D_2)\}, \quad (48b)$$

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, D_2, D_3), 1 - h(D_2), R_3'(D_3)\}, \quad (48c)$$

$$\{R_1, R_2, R_3\} = \{1, R_2'(D_2), R_3'(D_3)\}, \quad (48d)$$

where the arguments for independent and joint decoding of the auxiliary sources are similar to those presented in the other cases.

As in Section II-E5a, multiple cases can be considered in dependency of  $D_2$  and  $D_3$ . In the following we discuss two of those cases.

*Case 1*: Fig. 5 illustrates the inner and the outer bound for  $p_2 = 0.1$ ,  $p_3 = 0.2$ ,  $0 \leq D_i \leq D_{ic}, \forall i \in \{2, 3\}$ . In contrast to the case with two sources, we now define four rate regions and discuss the inner and the outer bound for each region simultaneously:

<sup>5</sup>First outer bound: lossless reconstruction of primary source; second outer bound: reconstruction of auxiliary source two with distortion constraint; third outer bound: reconstruction of auxiliary source three with distortion constraint; fourth outer bound: lossless reconstruction of primary source and reconstruction of auxiliary source two with distortion constraint; fifth outer bound: lossless reconstruction of primary source and reconstruction of auxiliary source three with distortion constraint; sixth outer bound: reconstruction of auxiliary source two and three with distortion constraints.

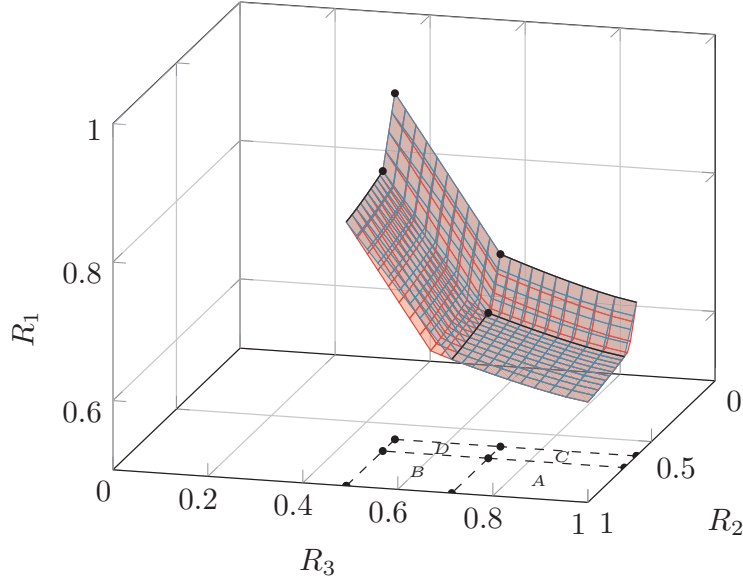


Fig. 5: The inner bound (blue area) and the outer bound (red area) for three binary sources, CI condition, and Hamming distortion measure with  $p_2 = 0.2$ ,  $p_3 = 0.3$ ,  $D_2 = 0.04$ ,  $D_3 = 0.05$ ,  $D_{2c} = 0.047$ , and  $D_{3c} = 0.145$ .

- 1) Region A for  $R_i \in (1 - h(D_i), 1], \forall i \in \{2, 3\}$ .
  - Outer bound: rate constraint set (43a),(43b) and (43c) maximizes the outer bound  $\mathcal{R}_c(D_2, D_3)$ . Both auxiliary sources are decoded without the primary source, i.e., in (43b), (43c) the rate constraint is  $R_i^*(d_i, \rho_i) = 1 - h(d_i)$  for  $0 \leq d_i \leq D_2, i = \{2, 3\}$ . The rate constraint for  $R_1$  is given in (43a). The outer bound (red area) for region A is illustrated in Fig. 5.
  - Inner bound: as for the outer bound, both auxiliary sources are decoded without the help of the primary source, (45a). The known achievable rate 3-tuple  $\{\xi(\cdot)\} = \{R_1, R_2, R_3\}$  is given in (46). The corresponding rate constraint for  $R_1$  is given in (44a). The inner bound (blue area) for region A is illustrated in Fig. 5.
  - The inner and the outer bound are tight for region A.
- 2) Region B for  $R_2 \in (1 - h(D_2), 1]$  and  $R_3 \in [R'_3(D_3), 1 - h(D_3)]$ .
  - Outer bound: either rate constraint set (43a), (43b), and (43c) or rate constraint set (43b), (43c), and (43e) maximizes the outer bound  $\mathcal{R}_c(D_2, D_3)$ . For the first rate constraint set, the auxiliary source three is decoded without the primary source and for the second rate constraint set the auxiliary source three is decoded with the primary source. Please note that in dependency on the decoding strategy for the auxiliary source three,  $R_3^*(d_3, \rho_3)$  in (43c) changes accordingly. The outer bound (red area) for region B is illustrated in Fig. 5.
  - Inner bound: the auxiliary source two is decoded without the primary source, as in (45a), and the auxiliary source three is decoded with the primary source, as in (45b), thus the two known achievable rate 3-tuple  $\{\xi(\cdot)\} = \{R_1, R_2, R_3\}$  are given in (47a) and (47b). Both known achievable rate 3-tuple for region B are illustrated in Fig. 5 as black lines. We can connect both lines with the time sharing argument forming the inner bound (blue area) for region B is illustrated in Fig. 5. The corresponding rate constraint for rate  $R_1$  is given in (44b).
  - The inner and the outer bound are not tight for region B, since for the outer bound the auxiliary source three is decoded without the primary source for some range of  $R_3$ .
- 3) Region C for  $R_2 \in [R'_2(D_2), 1 - h(D_2)]$  and  $R_3 \in (1 - h(D_3), 1]$  is similar to region B with index interchange.
- 4) Region D for  $R_2 \in [R'_2(D_2), 1 - h(D_2)]$  and  $R_3 \in [R'_3(D_3), 1 - h(D_3)]$ .
  - Outer bound: we distinguish between four rate constraint sets that can maximize the outer bound  $\mathcal{R}_c(D_2, D_3)$ .
    - a) (43a), (43b), and (43c): The auxiliary sources two and three are decoded without the primary source.
    - b) (43b), (43c), and (43d): The auxiliary source two is decoded with the primary source and the auxiliary source three is decoded without the primary source.
    - c) (43b), (43c), and (43e): The auxiliary source two is decoded without the primary source and the auxiliary source three is decoded with the primary source.
    - d) (43b), (43c), and (43f): The auxiliary sources two and three are decoded with the primary source.
 Please note that in dependency on the decoding strategy for the auxiliary sources,  $R_i^*(d_i, \rho_i)$  in (43b), (43c) changes accordingly. The outer bound (red area) for region D is illustrated in Fig. 5.
  - Inner bound: both auxiliary sources are decoded with the primary source, as in (45b), thus the four known achievable



rate 3-tuple  $\{\xi(\cdot)\} = \{R_1, R_2, R_3\}$  are given in (48a) - (48d). The four known achievable rate 3-tuple for region  $D$  are illustrated in Fig. 5 as black dots. We can connect all four dots with the time sharing argument. The inner bound (blue area) for region  $D$  is illustrated in Fig. 5. The corresponding rate constraint for rate  $R_1$  is given in (44d).

- The inner and the outer bound are not tight for region  $D$ , since for the outer bound the auxiliary sources two and three are decoded without the primary source for some range of  $R_2$  and  $R_3$ .

*Case 2:* Fig. 6 illustrates the inner and the outer bound for  $p_2 = 0.1$ ,  $p_3 = 0.2$  and  $D_{ic} < D_i, \forall i \in \{2, 3\}$ . With a larger distortion, the inner and the outer bound decrease, as expected. Apart from this, no other difference exists with respect to *Case 1*.

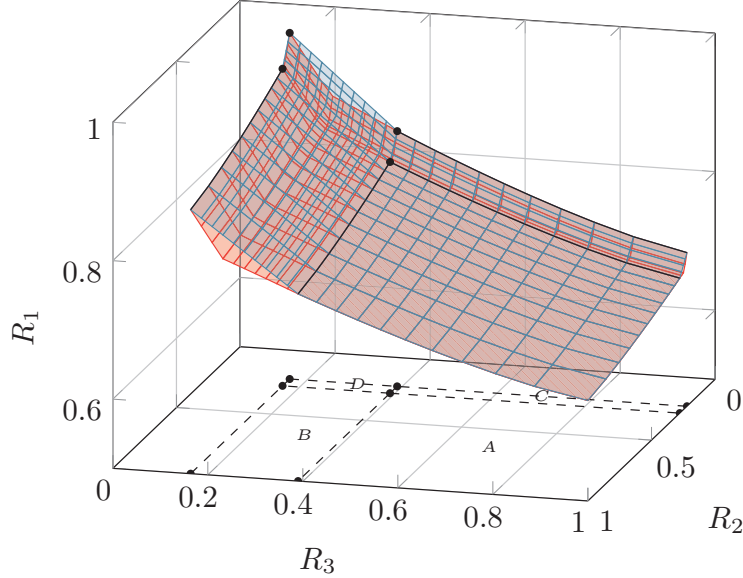


Fig. 6: The inner bound (blue area) and the outer bound (red area) for three binary sources, CI condition, and Hamming distortion measure with  $p_2 = 0.2$ ,  $p_3 = 0.3$ ,  $D_2 = 0.15$ ,  $D_3 = 0.25$ ,  $D_{2c} = 0.047$ , and  $D_{3c} = 0.145$ .

In conclusion, with  $p_2 < p_3$ , the inner and the outer bound are asymmetric, since  $R'_2(D_2) < R'_3(D_3)$ . In contrast to Section II-E5a, the inner and the outer bound do not coincide in general. However, the bounds are in general tight for the region  $R_i \in (1 - h(D_i), 1], \forall i \in \mathcal{L}$ . This holds for an arbitrary number of sources. This result is especially relevant for the weak many-help-one problem discussed in Section III.

We like to point out the impact of this result. If  $D_i = 0, \forall i \in \mathcal{L}$ , we end up with the Slepian-Wolf problem under the CI condition for  $N$  binary sources. In this case inner and outer bound coincide. To the best of our knowledge, no single-letter expression for this rate region has been published yet. Furthermore, we can show a continuous decrease of the inner and the outer bound with a continuous increase of  $D_i$  for any arbitrary cross-over probabilities  $p_i, \forall i \in \mathcal{L}$ . On the other hand, if  $D_i = p_i, \forall i \in \mathcal{L}$  which is the largest value possible for the distortion constraint, as pointed out in Section II-E1, we end up with the corresponding single-letter expressions for the weak many-help-one problem. Section III will discuss this result in detail.

### III. THE WEAK MANY-HELP-ONE PROBLEM

In this section, we investigate the weak many-help-one problem depicted in Fig. 1b, where the auxiliary sources work just as side information and thus are not reconstructed. The weak many-help-one problem is a special case of the discrete lossless CEO problem investigated by Gelfand and Pinsker [15]. Below we reproduce the Gelfand-Pinsker Theorem when specialized for the weak many-help-one problem.

[15, Theorem 1]: The admissible rate region  $\mathcal{R}$  is given by the set of all rate  $N$ -tuples  $R_{\mathcal{N}}$  such that there exists a  $N-1$ -tuple  $U_{\mathcal{L}}$  of discrete variables with pmf  $X_{\mathcal{N}}, U_{\mathcal{L}} \sim (5)$ , for which the following conditions are satisfied:

$$\sum_{i \in \mathcal{S}} R_i \geq \begin{cases} H(X_1 | U_{\mathcal{S}^c}) + I(X_{\mathcal{S} \setminus \{1\}}; U_{\mathcal{S} \setminus \{1\}} | X_1, U_{\mathcal{S}^c}) & \text{if } 1 \in \mathcal{S}, \\ I(X_{\mathcal{S}}; U_{\mathcal{S}} | X_1, U_{\mathcal{S}^c \setminus \{1\}}) & \text{otherwise,} \end{cases} \quad (49a)$$

$$(49b)$$

$\forall \mathcal{S} \subseteq \mathcal{N}$ , with  $\mathcal{S}^c$  being the complement of  $\mathcal{S}$ ; and for which there exists a function  $g_1(\cdot)$  such that

$$\mathbb{E}[d_1(X_1, g_1(X_1, U_{\mathcal{L}}))] \leq \epsilon. \quad (50)$$

### A. Independently Degraded Helpers

[15, Theorem 2]: If  $X_2, X_3, \dots, X_N$  are conditionally independent given  $X_1$ , i.e.,  $X_N \sim (4)$ , then the admissible rate region  $\mathcal{R}$  is the set of all rate tuples  $R_N$  such that there exists a tuple  $U_{\mathcal{L}}$  of discrete random variables with pmf  $X_N, U_{\mathcal{L}} \sim (13)$  for which the following conditions are satisfied:

$$\begin{cases} \sum_{i \in \mathcal{S}} R_i \geq H(X_1|U_{\mathcal{S}^c}) + \sum_{i \in \mathcal{S} \setminus \{1\}} I(X_i; U_i|X_1) & \text{if } 1 \in \mathcal{S}, \\ R_i \geq I(X_i; U_i|X_1) & \forall i \in \mathcal{L}, \end{cases} \quad (51a)$$

and for which there exists a function  $g_1(\cdot)$  such that

$$\mathbb{E}[d_1(X_1, g_1(X_1, U_{\mathcal{L}}))] \leq \epsilon. \quad (52)$$

### B. Binary sources and Hamming distortion

Based on [15, Theorem 2] and preliminary consideration in Section II-E, we can derive single-letter expressions for the weak many-help-one rate region with binary sources, CI condition, and Hamming distortion measure. This set up for the weak many-help-one problem is presented in Fig. 7. The properties of all random variables are identical to Section II-E. Single-letter expression for the mutual information  $I(X_i; U_i|X_1)$ ,  $\forall i \in \mathcal{L}$ , and for the conditional entropy  $H(X_1|U_{\mathcal{S}^c})$  in [15, Theorem 2] are found, in terms of the cross-over probabilities  $p_i$  and distortion constraints  $d_i$ ,  $\forall i \in \mathcal{L}$ . To derive  $\min I(X_i; U_i|X_1)$ , where the minimum is taken over all  $p(u_i|x_i)$ , and  $\min H(X_1|U_{\mathcal{S}^c})$ , where the minimum is taken over all  $p(u_{\mathcal{S}^c}|x_1)$  such that  $\mathbb{E}[d_H(X_1, g_1(X_1, U_{\mathcal{S}^c}))] \leq \epsilon$ , we can build on the same approach as in Section II-E.

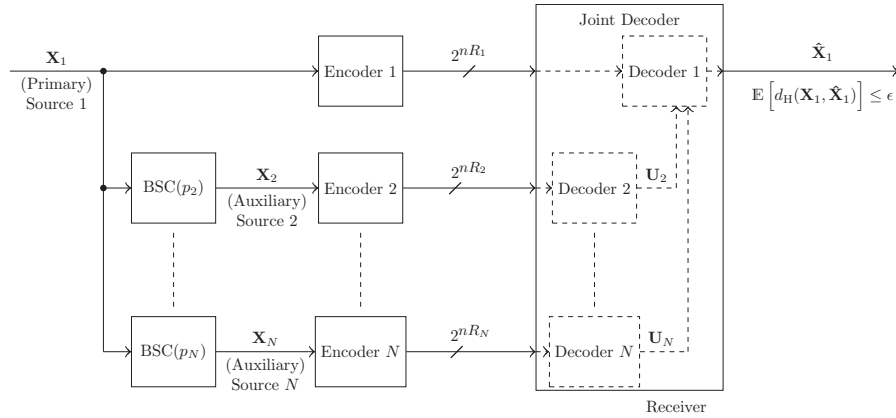


Fig. 7: The weak many-help-one problem for binary memory sources, CI condition, and Hamming distortion measure. Note that a particular decoding structure is shown in dashed lines, in which the auxiliary sources are decoded independently. As we shall demonstrate in Theorem 8, this structure is optimal for the weak many-help-one problem.

**Theorem 8:** If  $X_1$  takes values from a binary set with uniform probabilities,  $X_N \sim p(x_N) = p(x_1) \prod_{i=2}^N p(x_i|x_1)$ , binary random variables  $U_{\mathcal{L}}$  are distributed with pmf (13), then the admissible rate region  $\mathcal{R}$  is the set of all rate  $N$ -tuples  $R_N$  for which the following conditions are satisfied:

$$\begin{cases} R_1 \geq \phi(p_{\mathcal{L}}, d_{\mathcal{L}}), & (53a) \\ R_i \geq 1 - h(d_i), & (53b) \end{cases}$$

for  $0 \leq d_i \leq 0.5, \forall i \in \mathcal{L}$ .

*Proof:* The proof comprises three major arguments: (i) we show that for  $D_i = p_i, \forall i \in \mathcal{L}$ , the outer bounds of the weak and strong many-help-one problems coincide; (ii) for this particular case, we can conclude on the decoding strategy associated to the outer bound of the strong many-help-one problem; and (iii) from (i), it follows that this decoding strategy is optimal for the weak many-help-one problem, and thus the corresponding inner and outer bounds coincide.

- (i) Only the outer bound for perfect reconstruction of the primary source and no distortion constraints for the auxiliary sources maximizes the outer bound for the strong many-help-one problem in (31a) and (31b) if  $D_i = p_i, \forall i \in \mathcal{L}$ . Hence, the outer bounds of the strong and weak many-help-one problem are equivalent in this case.
- (ii) If  $D_i = p_i, \forall i \in \mathcal{L}$ , all sum rate constraints in (31a) are equal to or less than the rate constraint in (31a) for  $\mathcal{S} = \{1\}$ , i.e.,  $R_1 \geq H(X_1|U_{\mathcal{L}}) = \phi(p_{\mathcal{L}}, d_{\mathcal{L}})$ , implying independent decoding of all auxiliary sources, i.e.,  $R_i \geq R_i^* = I(X_i; U_i) = 1 - h(d_i)$ , for  $0 \leq d_i \leq 0.5, \forall i \in \mathcal{L}$  in (31b). Hence, the decoding strategy in Lemma 5 is optimal, i.e., all encoders operate independently and all auxiliary source decoders operate independently.

(iii) From [15, Theorem 2] we cannot unambiguously conclude on the decoding strategy of the auxiliary sources, eventually no auxiliary source decoding functions are defined and the mutual information in (51b) indicates a dependency on the primary source. However, with argument (i) we know that the decoding strategy in Lemma 5 is also optimal for the weak many-help-one outer bound, i.e., all sum rate constraints in (51a) are equivalent to or less than the rate constraint in (51a) for  $\mathcal{S} = \{1\}$ , i.e.,  $R_1 \geq H(X_1|U_{\mathcal{L}}) = \phi(p_{\mathcal{L}}, d_{\mathcal{L}})$ , implying independent decoding of all auxiliary sources, i.e.,  $R_i \geq I(X_i; U_i) = 1 - h(d_i)$ , for  $0 \leq d_i \leq 0.5, \forall i \in \mathcal{L}$  in (51b). As shown in Lemma 5, the inner and outer bound coincide for  $\min H(X_1|U_{\mathcal{L}})$  if  $X_1, U_{\mathcal{L}} \sim (24)$ . Since the weak many-help-one problem does not include specific distortion constraints  $D_i$  for the auxiliary sources, we can conclude that the inner and outer bound of the weak many-help-one problem coincide as well.

This completes the proof. ■

In summary, independent auxiliary source decoders are optimal for the weak many-help-one problem. This remarkable result is illustrated with the dashed blocks at the receiver side in Fig. 7. We cannot achieve any improvement by jointly decoding information provided by encoders  $i, \forall i \in \mathcal{L}$ .

Note that the strong many-help-one inner and outer bound do not coincide if  $D_i = p_i, \forall i \in \mathcal{L}$ .

1) *Discussion:* In this section we illustrate and discuss the results from Theorem 8 with two and three sources. In addition, for comparison, we include the inner and the outer bound for the strong many-help-one problem in Theorem 6 and Theorem 7, for  $D_2 = p_2$  and  $D_3 = p_3$ , and we point out similarities among the two problems.

a) *Two sources:* Fig. 8a illustrates the inner bound  $\mathcal{R}_a(p_2)$  and the outer bound  $\mathcal{R}_c(p_2)$  of the strong many-help-one problem for  $D_2 = p_2$ . The inner and the outer bound are given in (41a) - (42b) and (38a) - (38c), respectively.

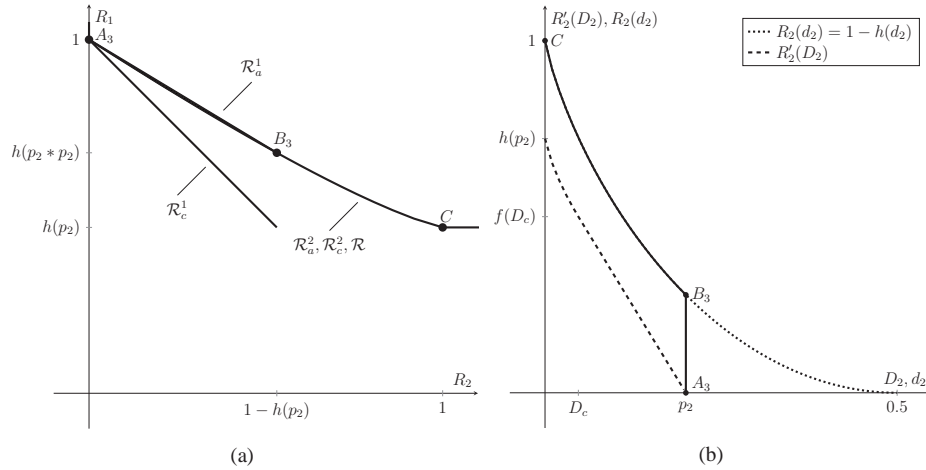


Fig. 8: (a) Rate region of weak many-help one problem  $\mathcal{R}$ , the inner bound  $\mathcal{R}_a = \max(\mathcal{R}_a^1, \mathcal{R}_a^2)$  and the outer bound  $\mathcal{R}_c = \mathcal{R}_c^2$  of strong many-help-one problem for  $D_2 = p_2$ , and (b) rate function for the second source under independent and joint decoding.

As pointed out in the proof of Theorem 8, the decoding strategy of Lemma 5 is optimal for the outer bound if  $D_2 = p_2$ , i.e., all encoders operate independently and the auxiliary source decoder operates independently. The outer bound of the strong many-help-one problem in Theorem 6 is thus given by

$$\begin{aligned} R_1 &\geq \phi(p_2, d_2) = h(p_2 * d_2), \\ R_2 &\geq 1 - h(d_2), \text{ for } 0 \leq d_2 \leq 0.5. \end{aligned}$$

The outer bound is illustrated in Fig. 8a. Note that the outer bound  $\mathcal{R}_c^2$  maximizes the outer bound  $\mathcal{R}_c(p_2)$  over the whole rate range of  $R_2$ , i.e.,  $\mathcal{R}_c(p_2) = \max[\mathcal{R}_c^1, \mathcal{R}_c^2] = \mathcal{R}_c^2$ .

For the inner bound  $\mathcal{R}_a(p_2)$ , point  $A_3$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{1, 0\}$ , and point  $B_3$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{h(p_2 * p_2), 1 - h(p_2)\}$ . Point  $A_3$  and  $B_3$  can be connected by the time-sharing argument. This is the inner bound  $\mathcal{R}_a^1$  for which it holds that the primary source is perfectly reconstructed and the auxiliary source is reconstructed with distortion constraint  $D_2 = p_2$ . In this case, the auxiliary source is jointly decoded. In Fig. 8b the rate  $R_2$  is illustrated. We see that point  $A_3$  is on the joint decoding rate function and point  $B_3$  is on the independent decoding rate function. An additional inner bound  $\mathcal{R}_a^2$  exists for which it holds that the primary source is perfectly reconstructed and the auxiliary source

is reconstructed with distortion constraint  $d_2$  for  $0 \leq d_2 \leq p_2$ . In this case, the auxiliary source is independently decoded, thus the achievable rate 2-tuples is  $\{R_1, R_2\} = \{h(p_2 * d_2), 1 - h(d_2)\}$ . Where the achievable rate  $R_1$  is given by Lemma 5. In Fig. 8a point  $C$  is the achievable rate 2-tuple  $\{R_1, R_2\} = \{h(p_2), 1\}$  for  $d_2 = 0$ . We can connect point  $B_3$  and  $C$  for  $p_2 \geq d_2 \geq 0$  accordingly. The corresponding rate  $R_2$  is illustrated in Fig. 8b, where point  $B_3$  and point  $C$  are on the independent decoding rate function.

In conclusion, for the rate range  $R_2 \in [0, 1 - h(p_2)]$  the inner and the outer bound are not tight but close. Note that for the inner bound the auxiliary source is decoded with the primary source, whereas for the outer bound the auxiliary source is decoded without the primary source. In the rate range  $R_2 \in [1 - h(p_2), 1]$  the inner and the outer bound are tight. The auxiliary source is then decoded without the primary source.

The rate region of the weak many-help-one  $\mathcal{R}$  is illustrated in Fig. 8a and it is identical to the outer bound of the strong many-help-one problem. Since the coding strategy for the strong many-help-one problem outer bound is decoding the auxiliary source without the primary source, we know that this is also true for the weak many-help-one problem as pointed out in the proof of Theorem 8.

*Remark:* The rate region of the weak many-help-one problem with two sources is equivalent to the result of Wyner [3] and Ahlswede and Körner [4] for binary source, CI condition, and Hamming distortion measure.

b) *Three sources:* Fig. 9 illustrates the inner bound  $\mathcal{R}_a(p_2, p_3)$  and the outer bound  $\mathcal{R}_c(p_2, p_3)$  of the strong many-help-one problem for binary sources, CI condition, and Hamming distortion measure. Similarly to Section III-B1, Paragraph b) we define four rate regions for the inner bound. The arguments are identical to Section III-B1, Paragraph b) and we now point out the coding strategies for the different rate regions of the inner bound  $\mathcal{R}_a(p_2, p_3)$ .

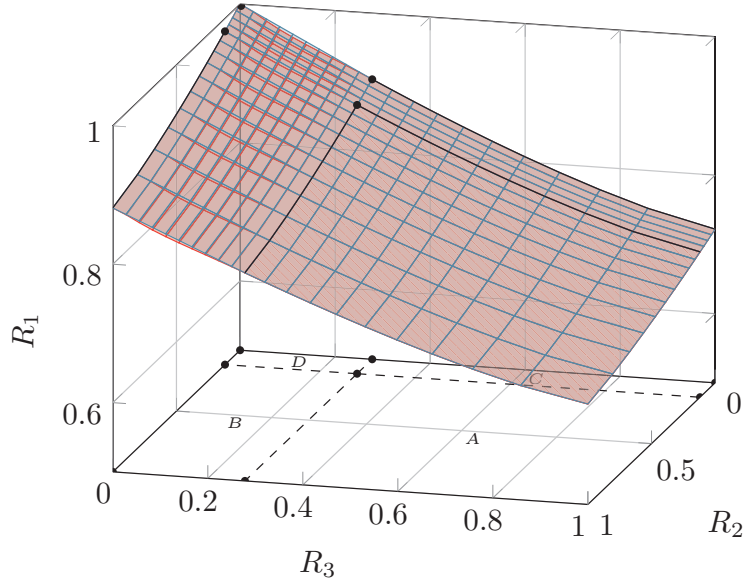


Fig. 9: The inner bound (blue area) and the outer bound (red area) and the admissible rate region (red area) for the strong and weak many-help-one problem, respectively, with  $p_2 = 0.2$ ,  $p_3 = 0.3$ .

- 1) Region A for  $R_i \in (1 - h(p_i), 1], i = 2, 3$ .

Both auxiliary sources are decoded without the primary source and the known achievable rate 3-tuple  $\{\xi(\cdot)\}$  is

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, d_2, d_3), 1 - h(d_2), 1 - h(d_3)\},$$

for  $0 \leq d_i \leq p_i, i = \{2, 3\}$ . The inner bound (blue area) for region A is illustrated in Fig. 9.

- 2) Region B for  $R_2 \in (1 - h(p_2), 1]$  and  $R_3 \in [0, 1 - h(p_3)]$ .

The auxiliary source two is decoded without the primary source and the auxiliary source three is decoded with the primary source, thus the two known achievable rate 3-tuples  $\{\xi(\cdot)\}$  are

$$\{R_1, R_2, R_3\} = \{\phi(p_2, p_3, d_2, p_3), 1 - h(d_2), 1 - h(p_3)\},$$

$$\{R_1, R_2, R_3\} = \{\phi(p_2, d_2), 1 - h(d_2), 0\},$$

for  $0 \leq d_2 < p_2$ . Both known achievable rate 3-tuples for region  $B$  are illustrated in Fig. 9 as black lines. We can connect both lines with the time sharing argument. The inner bound (blue area) for region  $B$  is illustrated in Fig. 9.

- 3) Region  $C$  for  $R_2 \in [0, 1 - h(p_2)]$  and  $R_3 \in (1 - h(p_3), 1]$  is similar to region  $B$  with index interchange.
- 4) Region  $D$  for  $R_2 \in [0, 1 - h(p_2)]$  and  $R_3 \in [0, 1 - h(p_3)]$ .

Both auxiliary sources are decoded with the primary source, thus the four known achievable rate 3-tuples  $\{\xi(\cdot)\}$  are

$$\begin{aligned}\{R_1, R_2, R_3\} &= \{\phi(p_2, p_3, p_2, p_3), 1 - h(p_2), 1 - h(p_3)\}, \\ \{R_1, R_2, R_3\} &= \{\phi(p_2, p_2), 1 - h(p_2), 0\}, \\ \{R_1, R_2, R_3\} &= \{\phi(p_3, p_3), 0, 1 - h(p_3)\}, \\ \{R_1, R_2, R_3\} &= \{1, 0, 0\}.\end{aligned}$$

The four known achievable rate 3-tuples for region  $D$  are illustrated in Fig. 9 as black dots. We can connect all four dots with the time sharing argument. The inner bound (blue area) for region  $D$  is illustrated in Fig. 9.

The outer bound  $\mathcal{R}_c(p_2, p_3)$  is illustrated in Fig. 9 as red area. As pointed out in the Proof of Theorem 8, the decoding strategy of Lemma 5 is optimal for the outer bound if  $D_2 = p_2$  and  $D_3 = p_3$ , i.e., all encoders operate independently and the auxiliary source decoders operate independently. The outer bound of the strong many-help-one problem in Theorem 6 is thus given by

$$\begin{aligned}R_1 &\geq \phi(p_2, p_2, d_2, d_3) = h(p_2 * d_2) + h(p_3 * d_3) - h(p_2 * d_2 * p_3 * d_3), \\ R_2 &\geq 1 - h(d_2), \text{ for } 0 \leq d_2 \leq 0.5, \\ R_3 &\geq 1 - h(d_3), \text{ for } 0 \leq d_3 \leq 0.5.\end{aligned}$$

In summary, the inner and the outer bound are tight for region  $A$ . For region  $B, C, D$  the inner and the outer bound are not tight but still close.

The rate region  $\mathcal{R}$  of the weak many-help-one problem is identical to the outer bound  $\mathcal{R}_c(p_2, p_3)$  of the strong many-help-one problem and illustrated in Fig. 9 as red area. As argued in the proof of Theorem 8, decoding the auxiliary source without the primary source is the optimal coding strategy for the strong many-help-problem outer bound, and this is also optimal for the weak many-help-one problem.

Note, with  $p_2 < p_3$ , the inner and the outer bound of the strong many-help-one problem and the admissible rate region of the weak many-help-one problem are asymmetric, since  $\phi(p_2, p_3, p_2, 0) < \phi(p_2, p_3, 0, p_3)$ .

In conclusion, we have derived the rate region of the weak many-help-one problem for binary sources, CI condition, and Hamming distortion measure. By comparison to the outer bound of the strong many-help-one problem, we have made an important statement on the optimal decoding strategy of the weak many-help-one problem. We have found that independent decoding of the auxiliary sources without any side information (neither primary source nor other auxiliary sources) is an optimal strategy as illustrated in Fig. 7. This is a remarkable result having operational meaning, as it allows for a significant reduction of the decoder complexity. In particular, it finds important application in emerging cooperative communications techniques based on lossy multi-route relaying links.

## APPENDIX A PROOF OF THEOREM 1

The proof of Theorem 1 is along the line of the proof for [13, Theorem 2] and [22, Theorem 1]. In [13] two sources are considered with a comprehensive proof and in [22] the system is extended to  $N$  sources. From now on we merely reference [13] for convenience. Here we point out the main difference and give additional analytics if necessary. The notation of [13, Theorem 2] and our notation correspond as follows:  $\mathbf{S}$  with  $\mathbf{X}$ ,  $\mathbf{W}$  with  $\mathbf{U}$ ,  $\mathbf{Z}$  with  $\emptyset$ , i.e., no side information  $\mathbf{Z}$  exists in our system, and  $\mathcal{T}$  with  $\mathcal{N}$ . The codebook generation is consistent except that encoder one labels the source sequence  $\mathbf{X}_1$  of length  $n$  with  $m_1 \in \{1, 2, \dots, 2^{nH(X_1)}\}$ , i.e.,  $R'_1$  in [13, Proof of Theorem 2] is equivalent to  $H(X_1)$  in order to meet the distortion constraint  $D_1 = 0$  for the primary source. To analyze the probability of error we give an additional lemma to include the event that an alternative choice of codewords exists which form a jointly typical tuple.

*Lemma 9:* If  $(\mathbf{X}_2, \dots, \mathbf{X}_N, \mathbf{Y}_2, \dots, \mathbf{Y}_M) \sim p(\mathbf{x}_2, \dots, \mathbf{x}_N) \cdot p(\mathbf{y}_2) \cdot \dots \cdot p(\mathbf{y}_M)$ , i.e., all  $\mathbf{Y}_m, m = 2, \dots, M$  are independent, then

$$\Pr \left[ (\mathbf{X}_2, \dots, \mathbf{X}_N, \mathbf{Y}_2, \dots, \mathbf{Y}_M) \in \mathcal{T}_\epsilon^{*(n)} \right] \leq 2^{n(H(X_2, \dots, X_N) + H(Y_2) + \dots + H(Y_M) - H(X_2, \dots, X_N, Y_2, \dots, Y_N) - \epsilon_2)} \quad (54)$$

where  $\epsilon_2 \rightarrow 0$  as  $\epsilon \rightarrow 0$  and  $n \rightarrow \infty$ .

*Proof:* The lemma can be established along the lines of the proof for [18, Lemma 10.6.2]. We here give the explicit argument for the case of weak typicality, i.e., extending the proof of [18, Theorem 7.6.1] for a index series of random time vectors,

$$\begin{aligned}
\Pr \left[ (\mathbf{X}_2, \dots, \mathbf{X}_N, \mathbf{Y}_2, \dots, \mathbf{Y}_M) \in \mathcal{T}_\epsilon^{(n)} \right] &= \sum_{(\mathbf{x}_2, \dots, \mathbf{x}_N, \mathbf{y}_2, \dots, \mathbf{y}_M) \in \mathcal{T}_\epsilon^{(n)}} p(\mathbf{x}_2, \dots, \mathbf{x}_N) \cdot p(\mathbf{y}_2) \cdot \dots \cdot p(\mathbf{y}_M) \\
&\leq 2^{n(H(X_2, \dots, X_N, Y_2, \dots, Y_M) + \epsilon)} \cdot 2^{-n(H(X_2, \dots, X_N) - \epsilon)} \\
&\quad \times 2^{-n(H(Y_2) - \epsilon)} \cdot \dots \cdot 2^{-n(H(Y_M) - \epsilon)} \\
&= 2^{-n(H(X_2, \dots, X_N) + H(Y_2) + \dots + H(Y_M) - H(X_2, \dots, X_N, Y_2, \dots, Y_M) - (M+1)\epsilon)} \quad (55)
\end{aligned}$$

where  $\mathcal{T}_\epsilon^{(n)}$  denotes the (weak) typical set as defined in [23, Chapter 2.5]. This completes the proof of Lemma 9.  $\blacksquare$

With Lemma 9 and [13, Proof of Theorem 2] we can analyze the probability of error. We obtain similar rate constraints for  $R'_i$  and  $R_i$ ,  $\forall i \in \mathcal{N}$  as in [13, Proof of Theorem 2, (40) - (42)]. Similar to [13, Proof of Theorem 2, (43) - (45)] we can derive

$$\begin{aligned}
\sum_{i \in \mathcal{S}} R_i &\geq H(X_1) + \sum_{i \in \mathcal{S} \setminus \{1\}} [I(X_i; U_i) - H(U_i)] - H(X_1) - H(U_{\mathcal{S}^c}) + H(X_1, U_{\mathcal{L}}) \\
&= \sum_{i \in \mathcal{S} \setminus \{1\}} [I(X_i, X_1, U_{\mathcal{S}^c}; U_i) - H(U_i)] - H(X_1, U_{\mathcal{S}^c}) + H(X_1 | U_{\mathcal{S}^c}) + H(X_1, U_{\mathcal{L}}) \quad (56)
\end{aligned}$$

$$= H(X_1 | U_{\mathcal{S}^c}) + \sum_{i \in \mathcal{S} \setminus \{1\}} [I(X_i, X_1, U_{\mathcal{S}^c}; U_i) - I(U_i; X_1, U_{\mathcal{S}^c})] - I(U_{\mathcal{S}_2}; U_{\mathcal{S}_3}; \dots; U_{\mathcal{S}_{|\mathcal{S} \setminus \{1\}|}} | X_1, U_{\mathcal{S}^c}) \quad (57)$$

$$= H(X_1 | U_{\mathcal{S}^c}) + \sum_{i \in \mathcal{S} \setminus \{1\}} I(X_i; U_i | X_1, U_{\mathcal{S}^c}) - I(U_{\mathcal{S}_2}; U_{\mathcal{S}_3}; \dots; U_{\mathcal{S}_{|\mathcal{S} \setminus \{1\}|}} | X_1, U_{\mathcal{S}^c}) \quad (58)$$

$$= H(X_1 | U_{\mathcal{S}^c}) + I(X_{\mathcal{S} \setminus \{1\}}; U_{\mathcal{S} \setminus \{1\}} | X_1, U_{\mathcal{S}^c}) \quad \text{if } 1 \in \mathcal{S}, \quad (59)$$

$$\begin{aligned}
\sum_{i \in \mathcal{S}} R_i &\geq \sum_{i \in \mathcal{S}} [I(X_i; U_i) - H(U_i)] - H(X_1, U_{\mathcal{S}^c \setminus \{1\}}) + H(X_1, U_{\mathcal{L}}) \\
&= \sum_{i \in \mathcal{S}} [I(X_i, X_1, U_{\mathcal{S}^c \setminus \{1\}}; U_i) - I(U_i; X_1, U_{\mathcal{S}^c \setminus \{1\}})] - I(U_{\mathcal{S}_2}; U_{\mathcal{S}_3}; \dots; U_{\mathcal{S}_{|\mathcal{S}|}} | X_1, U_{\mathcal{S}^c \setminus \{1\}}) \quad (60)
\end{aligned}$$

$$= \sum_{i \in \mathcal{S}} I(X_i; U_i | X_1, U_{\mathcal{S}^c \setminus \{1\}}) - I(U_{\mathcal{S}_2}; U_{\mathcal{S}_3}; \dots; U_{\mathcal{S}_{|\mathcal{S}|}} | X_1, U_{\mathcal{S}^c \setminus \{1\}}) \quad (61)$$

$$= I(X_{\mathcal{S}}; U_{\mathcal{S}} | X_1, U_{\mathcal{S}^c \setminus \{1\}}) \quad \text{if } 0 \notin \mathcal{S}, \quad (62)$$

$\forall \mathcal{S} \subseteq \mathcal{N}$ . The steps are justified as follows: in (56) we have used that  $U_i$  and  $X_1, U_{\mathcal{S}^c}$  are conditionally independent given  $X_i$ , and the chain rule of conditional entropy; some algebraic manipulations of mutual information and multivariate mutual information [20] omitted here yield (57); for (58) we applied the chain rule of mutual information; and (59) follows from the definition of the multivariate mutual information. Steps (60) - (62) can be justified just as before. This completes the proof of Theorem 1.  $\blacksquare$

## APPENDIX B PROOF OF THEOREM 2

The proof of Theorem 2 is along the line of the proof for [13, Theorem 3]. Here, we point out the main differences and give additional derivations where necessary. The notation of [13, Theorem 3] and our notation correspond as follows: see Appendix A. With the distortion constraint  $D_1 = 0$  we can rewrite the mutual information in [13, Proof of Theorem 3, (46) (b)] as follows:

$$I(\mathbf{X}_{\mathcal{N}}; T_{\mathcal{S}} | T_{\mathcal{S}^c}) = I(\mathbf{X}_1; T_{\mathcal{S}} | T_{\mathcal{S}^c}) + I(\mathbf{X}_{\mathcal{L}}; T_{\mathcal{S}} | \mathbf{X}_1, T_{\mathcal{S}^c}) \quad (63)$$

$$= H(\mathbf{X}_1 | T_{\mathcal{S}^c}) + I(\mathbf{X}_{\mathcal{L}}; T_{\mathcal{S}} | \mathbf{X}_1, T_{\mathcal{S}^c}) \quad (64)$$

where  $T_i$  is the index provided by encoder  $i$ ,  $\forall i \in \mathcal{N}$ . With the constraint of perfect reconstruction of  $\mathbf{X}_1$  we know that  $H(\mathbf{X}_1 | T_{\mathcal{N}}) = 0$  which justifies (64). Further reformulations are similar to (56) - (62) and [13, Proof of Theorem 3, (46) and (47)]. This completes the proof of Theorem 2.  $\blacksquare$



## APPENDIX C

### PROOF OF COROLLARY 3

The proof of Corollary 3 is along the line of the proof for [13, Corollary 4]. The notation of [13, Corollary 4] and our notation correspond as follows: see Appendix A. As proved in [13, Corollary 4, (16) and (17)] the mutual information in rate conditions (6a) and (6b) can be simplified to

$$I(X_{\mathcal{V}}; U_{\mathcal{V}} | X_1, U_{\mathcal{V}^c}) = \sum_{i=v_1}^{v_{|\mathcal{V}|}} I(X_{\mathcal{V}}; U_i | X_1, U_{\{v_1, \dots, i\}}, U_{\mathcal{V}^c}) \quad (65)$$

$$= \sum_{i=v_1}^{v_{|\mathcal{V}|}} I(X_i; U_i | X_1) \quad (66)$$

with  $\mathcal{V} = \mathcal{S} \setminus \{1\}, \mathcal{V}^c = \mathcal{S}^c$  for (6a) and  $\mathcal{V} = \mathcal{S}, \mathcal{V}^c = \mathcal{S}^c \setminus \{1\}$  for (6b). With (66) all sum rate bounds ( $|\mathcal{S}| > 1$ ) in (6b) become merely the sum of all single rate bounds ( $|\mathcal{S}| = 1$ ), hence all constraints in (6b) with  $|\mathcal{S}| > 1$  can be omitted. No simplification of the conditional entropy  $H(X_1 | U_{\mathcal{S}^c})$  in (6a) is possible. We know from Theorem 1 that if  $p(u_{\mathcal{L}}, x_{\mathcal{N}})$  satisfies (5), the distortion constraints  $D_i$  is achieved if there exists a function  $g_i(x_1, u_{\mathcal{L}}), \forall i \in \mathcal{L}$ . In [13, Proof of Corollary 5, (20)] it is shown that  $D_i$  is also achieved if there exists a function  $g_i(x_1, u_i), \forall i \in \mathcal{L}$  on the condition that  $\mathbf{X}_{\mathcal{N}} \sim (4)$ . However, we cannot reduce the arguments in the function  $g_1(x_1, u_{\mathcal{L}})$ . Therefore, we know from Theorem 1 that we can achieve distortion constraint (15) with  $g_1(x_1, u_{\mathcal{L}})$ . This completes the proof of Corollary 3. ■

## APPENDIX D

### PROOF OF COROLLARY 4

The proof is along the line of the proof for [13, Corollary 5]. The notation of [13, Corollary 5] and our notation correspond as follows: see Appendix A. Similar to the proof of Corollary 3 we can simplify the mutual information in rate conditions (10a) and (10b) by noting that the rate of encoder  $i, \forall i \in \mathcal{L}$  is at least the rate for the single-source Wyner-Ziv problem with side information  $X_{\mathcal{N} \setminus \{i\}}$  at the decoder [13, Proof of Corollary 5, (25)]. It is shown in [13, Proof of Corollary 5, (26) - (29)] that  $D_i$  is also achieved if there exists a function  $g_i(x_1, u_i), \forall i \in \mathcal{L}$ . The mutual information for all rate constraints in (10a) and (10b) can be rewritten with the same arguments as in (65) and (66). No simplification of the conditional entropy  $H(X_1 | U_{\mathcal{S}^c})$  in (10a) is possible. For this reason we cannot reduce the arguments in the function  $g_1(x_1, u_{\mathcal{L}})$ . Therefore, we know from Theorem 2 that we can achieve distortion constraint (19) with  $g_1(x_1, u_{\mathcal{L}})$ . This completes the proof of Corollary 4. ■

## REFERENCES

- [1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.
- [2] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 226–228, Mar 1975.
- [3] A. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 21, no. 3, pp. 294–300, May 1975.
- [4] R. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 21, no. 6, pp. 629–637, Nov 1975.
- [5] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan 1976.
- [6] T. Berger, "Multiterminal source coding," *Lectures presented at CISM Summer School on the Information Theory Approach to Communications*, vol. 229, pp. 171–231, 1977.
- [7] S.-Y. Tung, "Multiterminal source coding," Ph.D. dissertation, Cornell Univ., School of Elect. Eng., Ithaca, 1978.
- [8] T. Berger and R. W. Yeung, "Multiterminal source encoding with one distortion criterion," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 228–236, Mar 1989.
- [9] I. Csiszár and J. Körner, "Towards a general theory of source networks," *IEEE Trans. Inf. Theory*, vol. 26, no. 2, pp. 155–165, Mar 1980.
- [10] T. Han and K. Kobayashi, "A unified achievable rate region for a general class of multiterminal source coding systems," *IEEE Trans. Inf. Theory*, vol. 26, no. 3, pp. 277–288, May 1980.
- [11] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov 1997.
- [12] S. Tavildar, P. Viswanath, and A. B. Wagner, "The Gaussian many-help-one distributed source coding problem," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 564–581, Jan 2010.
- [13] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2762–2768, Nov 2004.
- [14] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 219–221, Mar 1979.
- [15] S. I. Gelfand and M. S. Pinsker, "Coding of sources on the basis of observations with incomplete information," *Problemy Peredachi Informatsii*, vol. 15, no. 2, pp. 45–57, 1979.
- [16] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577–2593, July 2005.
- [17] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.
- [18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.
- [19] A. Wyner, "A theorem on the entropy of certain binary sequences and applications: part ii," *IEEE Trans. Inf. Theory*, vol. 19, no. 6, pp. 769–772, Nov 1973.

- [20] W. McGill, "Multivariate information transmission," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 93–111, September 1954.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [22] M. Gastpar, "On Wyner-Ziv networks," in *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, vol. 1, Nov 2003, pp. 855–859.
- [23] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.